

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 1, 2004, 20:41:05 ; Search time 2748 Seconds  
(without alignments)  
9943.088 Million cell updates/sec

Title: US-10-030-390-2

Perfect score: 5497

Sequence: 1 gaattcgaataagtcattt.....ctagtcgaagtgatttaata 5497

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3267054 seqs, 2485319735 residues

Total number of hits satisfying chosen parameters: 6534108

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:
- 3: /cgn2\_6/ptodata/1/pubpna/US03\_NEW\_PUB.seq:
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:
- 10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:
- 11: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:
- 13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:
- 14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:
- 15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:
- 16: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:
- 17: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:
- 18: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:
- 19: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	5010.8	91.2	5770	9	US-09-838-718A-7
2	5001.4	91.0	5906	9	US-09-838-718A-6
3	4999	90.9	5870	9	US-09-838-718A-8
4	4993.4	90.8	5230	9	US-09-838-718A-5
C 5	1434.8	26.1	3765	9	US-09-070-927A-287
C 6	1109.4	20.2	10929	15	US-10-032-393-7
C 7	876.6	15.9	9412	15	US-10-032-393-51
C 8	876.6	15.9	9412	15	US-10-032-393-52
C 9	876.6	15.9	9417	15	US-10-032-393-48
C 10	876.6	15.9	9425	15	US-10-032-393-46
C 11	876.6	15.9	9437	15	US-10-032-393-17
C 12	876.6	15.9	9484	15	US-10-032-393-49
C 13	876.6	15.9	9566	15	US-10-032-393-50
C 14	876.6	15.9	12733	15	US-10-032-393-47

C 15	876.6	15.9	12739	15	US-10-032-393-8
C 16	833.8	15.2	9749	9	US-09-070-927A-154
C 17	747	13.6	747	9	US-09-974-300-5684
C 18	731.6	13.3	738	10	US-09-989-643-114
C 19	381	6.9	381	9	US-09-974-300-1224
C 20	244.2	4.4	6405	10	US-09-869-855A-1
C 21	244.2	4.4	6837	10	US-09-928-847B-49
C 22	241	4.4	735	10	US-09-989-643-115
C 23	229.2	4.2	732	10	US-09-989-643-113
C 24	204	3.7	447	16	US-10-166-349-11
C 25	204	3.7	447	16	US-10-166-349-19
C 26	178.8	3.3	1650	16	US-10-398-221-2127
C 27	178.8	3.3	1716	16	US-10-398-221-649
C 28	178.8	3.3	81905	16	US-10-398-221-2057
C 29	178.8	3.3	83689	16	US-10-398-221-11
C 30	174.6	3.2	6405	10	US-09-869-855A-1
C 31	174.6	3.2	6837	10	US-09-928-847B-49
C 32	143.2	2.6	5502	9	US-09-735-457-785
C 33	143.2	2.6	5502	9	US-09-902-941-785
C 34	143.2	2.6	5502	9	US-09-849-626-785
C 35	143.2	2.6	5502	10	US-09-476-300-785
C 36	143.2	2.6	5502	13	US-10-283-017-785
C 37	143.2	2.6	5502	15	US-10-017-754-785
C 38	143.2	2.6	5502	15	US-10-113-872-785
C 39	142.2	2.6	673	15	US-10-082-973-5
C 40	142.2	2.6	685	15	US-10-082-973-4
C 41	142.2	2.6	686	15	US-10-082-973-52
C 42	142	2.6	142	9	US-09-060-878-1
C 43	142	2.6	142	13	US-10-350-250-1
C 44	141.6	2.6	2290	10	US-09-935-280-16
C 45	140.8	2.6	4557	16	US-10-457-372-5

ALIGNMENTS

RESULT 1

US-09-838-718A-7  
; Sequence 7, Application US/09838718A  
; Patent No. US20020019043A1  
; GENERAL INFORMATION:  
; APPLICANT: Steidler, Lothar  
; APPLICANT: Remaut, Erik R.  
; APPLICANT: Fiers, Walter R.  
; TITLE OF INVENTION: USE OF A CYTOKINE-PRODUCING LACTOCOCCUS STRAIN TO TREAT COLITIS  
; FILE REFERENCE: 2676-4779US  
; CURRENT APPLICATION NUMBER: US/09/838,718A  
; CURRENT FILING DATE: 2001-04-19  
; PRIOR APPLICATION NUMBER: PCT/EP99/07800  
; PRIOR FILING DATE: 1999-10-06  
; PRIOR APPLICATION NUMBER: EP 98203529.7  
; PRIOR FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 7  
; LENGTH: 5770  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: plasmid pT1MIL10  
US-09-838-718A-7

Query Match	91.2%;	Score 5010.8;	DB 9;	Length 5770;	
Best Local Similarity	100.0%;	Pred. No. 0;			
Matches 5002;	Conservative	0;	Mismatches	0;	
		Indels	0;	Gaps	0;
QY	496	CTAACTAGTAGATCGGCTGTCTAAACAAAGCCCGAAAGAGCTGAGTTGGTCTCTGCCAC	555		
Db	769	CTAACTAGTAGATCGGCTGTCTAAACAAAGCCCGAAAGAGCTGAGTTGGTCTCTGCCAC	828		
QY	556	CGCTGAGCAATAACTAGCATAACCCCTTGGGCGCTCTAAACGGGTCTTGAGGGGTTTTT	615		
Db	829	CGCTGAGCAATAACTAGCATAACCCCTTGGGCGCTCTAAACGGGTCTTGAGGGGTTTTT	888		

Qy	616	GCTGAAAGGAGGAAC	TATATCCGATGACCTGCAGGCAAGCTCTAGAAATCGATACGATTT	675
Db	889	GCTGAAAGGAGGAAC	TATATCCGATGACCTGCAGGCAAGCTCTAGAAATCGATACGATTT	948
Qy	676	TGAAGTGGCAACAGAT	AAAAAAGCAGTTTAAATTTGTTGCTGAACTTTTAAACCAAGC	735
Db	949	TGAAGTGGCAACAGAT	AAAAAAGCAGTTTAAATTTGTTGCTGAACTTTTAAACCAAGC	1008
Qy	736	AAATACAATCATTTGT	CCCAACAGATACGACAGAGAGGCGAAAAATTCCTCCCTGGTCGAT	795
Db	1009	AAATACAATCATTTGT	CCCAACAGATACGACAGAGAGGCGAAAAATTCCTCCCTGGTCGAT	1068
Qy	796	CATTATAAAGCAAAATGCC	TTTCTAAAGATATAAAGCTATAAAGACTATGGATCATATAG	855
Db	1069	CATTATAAAGCAAAATGCC	TTTCTAAAGATATAAAGCTATAAAGACTATGGATCATATAG	1128
Qy	856	TTTAGAAAAAGATGTGAT	CCGTACGCGTTTCAAAATTTGCAACCCAGGAATGAATTACTA	915
Db	1129	TTTAGAAAAAGATGTGAT	CCGTACGCGTTTCAAAATTTGCAACCCAGGAATGAATTACTA	1188
Qy	916	TCCTTTTTCAGAGCGG	CAAAAGAAAAAGCAATGATACACCAATCAGTCGCAAAAAA	975
Db	1189	TCCTTTTTCAGAGCGG	CAAAAGAAAAAGCAATGATACACCAATCAGTCGCAAAAAA	1248
Qy	976	GATATAATGGAGATAAG	ACGGTTTCGTGTTTCGTCTGACTTGCAACCATATCATATAAAATC	1035
Db	1249	GATATAATGGAGATAAG	ACGGTTTCGTGTTTCGTCTGACTTGCAACCATATCATATAAAATC	1308
Qy	1036	GAACACAGCAAGAAATG	CGGGAACGTTAAAGAGTTATGAAATAGACATTAGNAGCAA	1095
Db	1309	GAACACAGCAAGAAATG	CGGGAACGTTAAAGAGTTATGAAATAGACATTAGNAGCAA	1368
Qy	1096	CTTAAGAGTGTGTTGAT	GTAGTGCAGTATCTTAAATTTTGTATATAGCAATTCGAAGTTAA	1155
Db	1369	CTTAAGAGTGTGTTGAT	GTAGTGCAGTATCTTAAATTTTGTATATAGCAATTCGAAGTTAA	1428
Qy	1156	ATTAGATGCTAAAAATTT	TGTAATTAAGAGAGTGATTAATGAACAAAAATATAAATA	1215
Db	1429	ATTAGATGCTAAAAATTT	TGTAATTAAGAGAGTGATTAATGAACAAAAATATAAATA	1488
Qy	1216	TTCTCAAAACCTTTT	TAAACGAGTGAAGTACTCAACCAATATAAAAACAATTTGAATTT	1275
Db	1489	TTCTCAAAACCTTTT	TAAACGAGTGAAGTACTCAACCAATATAAAAACAATTTGAATTT	1548
Qy	1276	AAAAAGAACCGGATAC	CGGTTTACGAATTTGGAAACAGGTTAAAGGGGATTTACGACGAAACT	1335
Db	1549	AAAAAGAACCGGATAC	CGGTTTACGAATTTGGAAACAGGTTAAAGGGGATTTACGACGAAACT	1608
Qy	1336	GGCTTAAATTAAGTAAA	ACAGGTAAACGTTATTAAGATAGACAGTCACTATTTCAACTTATC	1395
Db	1609	GGCTTAAATTAAGTAAA	ACAGGTAAACGTTATTAAGATAGACAGTCACTATTTCAACTTATC	1668
Qy	1396	GTCAGAAAAATTA	AAACTGGAATCTCGTGCACTTTTAATTCACCAAGATATCTACAGTT	1455
Db	1669	GTCAGAAAAATTA	AAACTGGAATCTCGTGCACTTTTAATTCACCAAGATATCTACAGTT	1728
Qy	1456	TCATTTCCCTTAAACAA	CAGAGGTATAAAATTTGTTGGGAGTATTCCTTTACCAATTTAAGCAC	1515
Db	1729	TCATTTCCCTTAAACAA	CAGAGGTATAAAATTTGTTGGGAGTATTCCTTTACCAATTTAAGCAC	1788
Qy	1516	ACAAATTTTAAAAAGT	GGTTTTTGAAGCCATCGCTGTCGACATCTATCTGATTTGTGA	1575
Db	1789	ACAAATTTTAAAAAGT	GGTTTTTGAAGCCATCGCTGTCGACATCTATCTGATTTGTGA	1848
Qy	1576	AGAAGGATTTCTACA	AGCGTACTCTGGATATTCACCGAACACTAGGGTTGCTCTTCGCACAC	1635
Db	1849	AGAAGGATTTCTACA	AGCGTACTCTGGATATTCACCGAACACTAGGGTTGCTCTTCGCACAC	1908
Qy	1636	TCAGTCTCGATTCAG	CAATTTGCTTCAGCGGAATGCTTTTCACTTAAACCAA	1695
Db	1909	TCAGTCTCGATTCAG	CAATTTGCTTCAGCGGAATGCTTTTCACTTAAACCAA	1968

[illegible]

|||||  
3049 ATGGAAGAGAAAAGCAATGACTCGTGTGTAATATGACGAATCGTTGCTTATTT 3108  
Qy TTTTAAAGCGGTATCTAGATATACGAAACAAACGAACCTGAATAGAAACGAAAAGA 2895  
Db TTTTAAAGCGGTATCTAGATATACGAAACAAACGAACCTGAATAGAAACGAAAAGA 3168  
Qy GCCATGACACATTTATAAAATGTTTGACGACATTTTATAAATGCATAGCCCGATAAGATT 2955  
Db GCCATGACACATTTATAAAATGTTTGACGACATTTTATAAATGCATAGCCCGATAAGATT 3228  
Qy GCCAAACCAACGGCTATACGTTAGTCAGATGAACCTCTCCCTCGTAAAGATTATTTAAT 3015  
Db GCCAAACCAACGGCTATACGTTAGTCAGATGAACCTCTCCCTCGTAAAGATTATTTAAT 3288  
Qy TAACTTTGTTTGAAGCGGTATATACCGTACTATCATTTATATAGGGAATCAGAGATT 3075  
Db TAACTTTGTTTGAAGCGGTATATACCGTACTATCATTTATATAGGGAATCAGAGATT 3348  
Qy TTCAAGTATCTAAGCTACTGAATTTAAGAAATGTTAAGCAATCAATCGGAAATCGTTGA 3135  
Db TTCAAGTATCTAAGCTACTGAATTTAAGAAATGTTAAGCAATCAATCGGAAATCGTTGA 3408  
Qy TTGCTTTTTCGTTATTCATTTATAGAGTGGAGTTGTGATGAATCATGATGAATGTAAA 3195  
Db TTGCTTTTTCGTTATTCATTTATAGAGTGGAGTTGTGATGAATCATGATGAATGTAAA 3468  
Qy ACTTATAAAAAATAGTTTAAATGGAGATAAGAAATTAGCAATATCTATACACTAGAA 3255  
Db ACTTATAAAAAATAGTTTAAATGGAGATAAGAAATTAGCAATATCTATACACTAGAA 3528  
Qy ACGTTTAAAGAAAGTTAGAAAAGAGAAATATCTACTTAGAAAACAAAATCAGATAAGTAT 3315  
Db ACGTTTAAAGAAAGTTAGAAAAGAGAAATATCTACTTAGAAAACAAAATCAGATAAGTAT 3588  
Qy TTTTCTTCGAGGGGAGAGTTATATATATAAGTTAATAGAAATACAAAATTAATTTAT 3375  
Db TTTTCTTCGAGGGGAGAGTTATATATATAAGTTAATAGAAATACAAAATTAATTTAT 3648  
Qy TCGATTAGTGGAAAAAATGACTTATAAAGGAAAAAATCTTTTCAAACATGCAATA 3435  
Db TCGATTAGTGGAAAAAATGACTTATAAAGGAAAAAATCTTTTCAAACATGCAATA 3708  
Qy TTGAAACAGTTGTAATGAAAGCAACCAAGTTAATTAACAACCTATTTTATAGATTT 3495  
Db TTGAAACAGTTGTAATGAAAGCAACCAAGTTAATTAACAACCTATTTTATAGATTT 3768  
Qy ATAGGAAAGGAGAACAGCTGAATGAATATCCCTTTTGTGTAGAAACTGTGCTTCATGAC 3555  
Db ATAGGAAAGGAGAACAGCTGAATGAATATCCCTTTTGTGTAGAAACTGTGCTTCATGAC 3828  
Qy GGCTTGTAAAGTACAAATTTAAATAGTAAATTTGCTCAATCTACCAAGCCAGGT 3615  
Db GGCTTGTAAAGTACAAATTTAAATAGTAAATTTGCTCAATCTACCAAGCCAGGT 3888  
Qy AAAAGCAAAAGGGCTATTTTGGCTATCGCTCAAAATCAAGCATGATGGCGGTGCTGGT 3675  
Db AAAAGCAAAAGGGCTATTTTGGCTATCGCTCAAAATCAAGCATGATGGCGGTGCTGGT 3948  
Qy GTTGTCTGACTTCGAGGAAGCGATTCAAGAAAAATCAAGATACATTTACACATTGGACA 3735  
Db GTTGTCTGACTTCGAGGAAGCGATTCAAGAAAAATCAAGATACATTTACACATTGGACA 4008  
Qy CCCAACGTTTATCGTTATGAAACGTATGCAGAGCAAAACCGTTCAATACAGAAAGCAT 3795  
Db CCCAACGTTTATCGTTATGAAACGTATGCAGAGCAAAACCGTTCAATACAGAAAGCAT 4068  
Qy TCTGAAACAAATTTAAGCAAAATCAATACCTCTTTTATGATTTGATATTTCAACGGCA 3855  
Db TCTGAAACAAATTTAAGCAAAATCAATACCTCTTTTATGATTTGATATTTCAACGGCA 4128  
Qy AAAGAAACATTTTCAGCAAGCGATATTTTAAACACCGCTATTGATTTAGTTTATGCTT 3915

Db 4129 AAAGAAACATTTTTCAGCAAGCGATATTTTAAACACCGCTATTGATTTAGTTTATGCTT 4188  
Qy 3916 ACTATGATATCAAAATCTGATATAAGGTTTATCAAGCATATTTTGTGTTTGAACCGCATC 3975  
Db 4189 ACTATGATATCAAAATCTGATATAAGGTTTATCAAGCATATTTTGTGTTTGAACCGCATC 4248  
Qy 3976 TATGTGACTTCAAAATCAGAAATTTAAATCTCTGTAAGCAGCAAAATTAATTCGCAAAAT 4035  
Db 4249 TATGTGACTTCAAAATCAGAAATTTAAATCTCTGTAAGCAGCAAAATTAATTCGCAAAAT 4308  
Qy 4036 ATCCGAGAAATTTTGGAAAGTCTTTGCGAGTTGATCTAACGTGTATCATTTTGGTATT 4095  
Db 4309 ATCCGAGAAATTTTGGAAAGTCTTTGCGAGTTGATCTAACGTGTATCATTTTGGTATT 4368  
Qy 4096 GCTCCATACCAAGAAACGACATGTAGAAATTTTGTGATCTTAATACCGTTATCTTTC 4155  
Db 4369 GCTCCATACCAAGAAACGACATGTAGAAATTTTGTGATCTTAATACCGTTATCTTTC 4428  
Qy 4156 AAAGAAATGCGAAGATGCTCTTTCAAACACAGATAAAGGGCTTTACTCGTTCAAGT 4215  
Db 4429 AAAGAAATGCGAAGATGCTCTTTCAAACACAGATAAAGGGCTTTACTCGTTCAAGT 4488  
Qy 4216 CTAAACGGTTTAAAGCGGTACAGAGGCAAAACAAAGTAGATGACCTGTGTTAATCTC 4275  
Db 4489 CTAAACGGTTTAAAGCGGTACAGAGGCAAAACAAAGTAGATGACCTGTGTTAATCTC 4548  
Qy 4276 TTATTTGCAGAAACGAAATTTTTCAGGAGAAAAGGGTTTAATAGGCGTAAATACGTCATG 4335  
Db 4549 TTATTTGCAGAAACGAAATTTTTCAGGAGAAAAGGGTTTAATAGGCGTAAATACGTCATG 4608  
Qy 4336 TTTACCCCTCTCTTTAGCCCTACTCTTAGTTTACAGCTATTTCAATCGAAACGTCGGAATATAT 4395  
Db 4609 TTTACCCCTCTCTTTAGCCCTACTCTTAGTTTACAGCTATTTCAATCGAAACGTCGGAATATAT 4668  
Qy 4396 ATGTTTGAGTTTAAATAATCGATTAGATCAACCCCTTAGAAGAAAAGAAAGTAAATCAAAAT 4455  
Db 4669 ATGTTTGAGTTTAAATAATCGATTAGATCAACCCCTTAGAAGAAAAGAAAGTAAATCAAAAT 4728  
Qy 4456 GTTAGAAGTGCCTATTCAGAAAACTATCAAGGGGCTAATAGGGAATAATACGATTCCTT 4515  
Db 4729 GTTAGAAGTGCCTATTCAGAAAACTATCAAGGGGCTAATAGGGAATAATACGATTCCTT 4788  
Qy 4516 TSCAAAGCTTGGGTATCAAGTGATTTAAACAGTAAAGATTTATTTGTCGGTCAAGGGTGG 4575  
Db 4789 TSCAAAGCTTGGGTATCAAGTGATTTAAACAGTAAAGATTTATTTGTCGGTCAAGGGTGG 4848  
Qy 4576 TTTAAATTTCAAGAAAAAAGAAAGCGACGTCAACGTTCTTCTCAGAAATGGAAGAA 4635  
Db 4849 TTTAAATTTCAAGAAAAAAGAAAGCGACGTCAACGTTCTTCTCAGAAATGGAAGAA 4908  
Qy 4636 GATTTTAAAGTGTATTTAGCGAAGAAAGCGATGTATACAGCCCTTATTTAGTGACGACC 4695  
Db 4909 GATTTTAAAGTGTATTTAGCGAAGAAAGCGATGTATACAGCCCTTATTTAGTGACGACC 4968  
Qy 4696 AAAAAAGAGATTAGAGAAAGTCTAGGCATTTCTGAAACGCAATTAGATAAAATTTGCTGAAG 4755  
Db 4969 AAAAAAGAGATTAGAGAAAGTCTAGGCATTTCTGAAACGCAATTAGATAAAATTTGCTGAAG 5028  
Qy 4756 GTACTGAAGGGCAATCAGAAATTTTCTTAAAGATTAAACAGGAAGAAATTTGTTGGCATT 4815  
Db 5029 GTACTGAAGGGCAATCAGAAATTTTCTTAAAGATTAAACAGGAAGAAATTTGTTGGCATT 5088  
Qy 4816 CAACTTGTCTAGTTTAAATCAATTTGTGCTATCGATCATTAAAGTAAAAAAGAAAGAAAA 4875  
Db 5089 CAACTTGTCTAGTTTAAATCAATTTGTGCTATCGATCATTAAAGTAAAAAAGAAAGAAAA 5148  
Qy 4876 GAAAGCTATATAAAGCGCTGACAAATTTCTTTGATTTAGAGCATACATTCATTCAAGAG 4935  
Db 5149 GAAAGCTATATAAAGCGCTGACAAATTTCTTTGATTTAGAGCATACATTCATTCAAGAG 5208  
Qy 4936 ACTTTAAACAGCTAGCAGAACCGCTTAAACAGGACACACAACTCGATTTGTTTATGCTAT 4995  
Db 5209 ACTTTAAACAGCTAGCAGAACCGCTTAAACAGGACACACAACTCGATTTGTTTATGCTAT 5268

QY	4996	GATACAGGCTGAAAAATAAAACCGCACATATGCCATTACATTTATATCTATGATACGTTGTT	5055
Db	5269	GATACAGGCTGAAAAATAAAACCGCACATATGCCATTACATTTATATCTATGATACGTTGTT	5328
QY	5056	TGTTTTTTCTTGTCTTTAGCGAATGATATAGCAGAAATATACAGAGTAAGATTTAAAT	5115
Db	5329	TGTTTTTTCTTGTCTTTAGCGAATGATATAGCAGAAATATACAGAGTAAGATTTAAAT	5388
QY	5116	AATTTATAGGGGAGAGAGAGAGTAGCCGAAACCTTTTGTGGCTTCGACTGAAACG	5175
Db	5389	AATTTATAGGGGAGAGAGAGAGTAGCCGAAACCTTTTGTGGCTTCGACTGAAACG	5448
QY	5176	AAGTGAGGAAAGGCTPACTAAACCGTCGAGGGCGAGTCAGAGCGAAGCGAACTTGAT	5235
Db	5449	AAGTGAGGAAAGGCTPACTAAACCGTCGAGGGCGAGTCAGAGCGAAGCGAACTTGAT	5508
QY	5236	TTTTTAATTTCTATCTTTTATAGGTCATTAGAGTATCTTATTTGTCTTATAACTATTT	5295
Db	5509	TTTTTAATTTCTATCTTTTATAGGTCATTAGAGTATCTTATTTGTCTTATAACTATTT	5568
QY	5296	AGCAGCATAATAGATTTATGTAATAGGTCATTAAAGTTGAGCATATTAGAGGAGGAAAT	5355
Db	5569	AGCAGCATAATAGATTTATGTAATAGGTCATTAAAGTTGAGCATATTAGAGGAGGAAAT	5628
QY	5356	CTTGAGAAATATTGAGAAACCGATATCATGGAATGGATAGTTCTTGTGGTTACGTG	5415
Db	5629	CTTGAGAAATATTGAGAAACCGATATCATGGAATGGATAGTTCTTGTGGTTACGTG	5688
QY	5416	GTTTTTAACTAAAGTAGTGAATTTTGTATTTTGGTGTGTGTCTTGTGTGTGTGTGT	5475
Db	5689	GTTTTTAACTAAAGTAGTGAATTTTGTATTTTGGTGTGTGTCTTGTGTGTGTGTGT	5748
QY	5476	TGCTAGTCAAAGTGATTAATA 5497	
Db	5749	TGCTAGTCAAAGTGATTAATA 5770	

RESULT 2

US-09-838-718A-6

; Sequence 6, Application US/09838718A

; Patent No. US20020019043A1

; GENERAL INFORMATION:

; APPLICANT: Steidler, Lothar

; APPLICANT: Remaut, Erik R.

; APPLICANT: Fiers, Walter R.

; TITLE OF INVENTION: USE OF A CYTOKINE-PRODUCING LACTOCOCCUS STRAIN TO TREAT COLITIS

; FILE REFERENCE: 2676-4779US

; CURRENT APPLICATION NUMBER: US/09/838,718A

; CURRENT FILING DATE: 2001-04-19

; PRIOR APPLICATION NUMBER: PCT/EP99/07800

; PRIOR FILING DATE: 1999-10-06

; PRIOR APPLICATION NUMBER: EP 98203529.7

; PRIOR FILING DATE: 1998-10-20

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 6

; LENGTH: 5906

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: plasmid pTINX

US-09-838-718A-6

Query Match 91.0%; Score 5001.4; DB 9; Length 5906;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 5002; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 495 TCTAATAGTAGATCCGGCTGTCTAACAAGCCCGAAGAGAGCTGAGTTGCTCTGCGCA 554

Db 904 TATAAGTAGTAGATCCGGCTGTCTAACAAGCCCGAAGAGAGCTGAGTTGCTCTGCGCA 963

QY 555 CCCTGAGCAATTAATAGCATTAACCCCTTGGGGCTCTTAACCGGCTCTTGAGGGGTTTTT 614



Db 2044 CTCAAGTCTCGATTTCAGCAATTCGTTAAGCTGCCAGGGAATGCTTTTCATCCTTAAACCAA 2103  
Qy 1695 AAGTAAACAGTGTCTTAAATAAACTTACCCGCCATACACAGATGTTCCAGATAAATATT 1754  
Db 2104 AAGTAAACAGTGTCTTAAATAAACTTACCCGCCATACACAGATGTTCCAGATAAATATT 2163  
Qy 1755 GGAAGCTATATACGTACTTGTGTTTCAAAATGGGTCAATCGAGAAATACGTCAACTGTTTA 1814  
Db 2164 GGAAGCTATATACGTACTTGTGTTTCAAAATGGGTCAATCGAGAAATACGTCAACTGTTTA 2223  
Qy 1815 CTAAATATCAGTTTCTACAGCAATGAAACACGCCAAAGTAAACAATTTAAAGTACCGTTA 1874  
Db 2224 CTAAATATCAGTTTCTACAGCAATGAAACACGCCAAAGTAAACAATTTAAAGTACCGTTA 2283  
Qy 1875 CTTATGAGCAAGTATGCTTATTTTAAATAGTTTATCTATTTATTTAAACGGGAGAAATAT 1934  
Db 2284 CTTATGAGCAAGTATGCTTATTTTAAATAGTTTATCTATTTATTTAAACGGGAGAAATAT 2343  
Qy 1935 TCTATGAGTCCGTTTGTAAATTTGGAAGTTACAGTTTACTTAAGGGGAACTGAGTAA 1994  
Db 2344 TCTATGAGTCCGTTTGTAAATTTGGAAGTTTACAGTTTACTTAAGGGGAACTGAGTAA 2403  
Qy 1995 TTATTAGGTATATCTACTGACAGCTTCCAAGGAGCTAAAGAGTCCCTAGCGCTCTTATCA 2054  
Db 2404 TTATTAGGTATATCTACTGACAGCTTCCAAGGAGCTAAAGAGTCCCTAGCGCTCTTATCA 2463  
Qy 2055 TGGGGAAGCTCGGATCATATGCAAGCAAAATAAACTCGCAACAGCACTTGGAGAAATGG 2114  
Db 2464 TGGGGAAGCTCGGATCATATGCAAGCAAAATAAACTCGCAACAGCACTTGGAGAAATGG 2523  
Qy 2115 GACGAATCGAGAAACCCCTTTTACGCTGGATTACATATCTAATAAGCGGTAAAGAGAC 2174  
Db 2524 GACGAATCGAGAAACCCCTTTTACGCTGGATTACATATCTAATAAGCGGTAAAGAGAC 2583  
Qy 2175 GGGTTCAAAAAGTTTAAATAAAGGAGAGCAATCAATGCAATGAGTAAAGTATATTTT 2234  
Db 2584 GGGTTCAAAAAGTTTAAATAAAGGAGAGCAATCAATGCAATGAGTAAAGTATATTTT 2643  
Qy 2235 TTGGACACCTGGAGAAATTTAGAACGCTCTCTCAAGACAGTTTACAAAGAGCTAGTG 2294  
Db 2644 TTGGACACCTGGAGAAATTTAGAACGCTCTCTCAAGACAGTTTACAAAGAGCTAGTG 2703  
Qy 2295 CACTAAACATAATTTAAACGCTATTAAGTGTGTGGAACTGCTATATATATGGAAGAGCCG 2354  
Db 2704 CACTAAACATAATTTAAACGCTATTAAGTGTGTGGAACTGCTATATATATGGAAGAGCCG 2763  
Qy 2355 TAGAAGAAATTAAGCAAGAGAGAAATTTAGAGAAATTTAATGCAATATGCGTGGCCGT 2414  
Db 2764 TAGAAGAAATTAAGCAAGAGAGAAATTTAGAGAAATTTAATGCAATATGCGTGGCCGT 2823  
Qy 2415 TAGATGGGAACATATCAATTTTCTGGAGAAATCAAAATTTGAAGATTTACATGACACTG 2474  
Db 2824 TAGATGGGAACATATCAATTTTCTGGAGAAATCAAAATTTGAAGATTTACATGACACTG 2883  
Qy 2475 GGCAAATGAATTTACGCTCTTTACGTATAAAGAGCGGTTTATCTTAAATAACGGCT 2534  
Db 2884 GGCAAATGAATTTACGCTCTTTACGTATAAAGAGCGGTTTATCTTAAATAACGGCT 2943  
Qy 2535 CTTTATTAAGAAATTCCTTAGCGTGTGTTTCTCGAAATGCTCGCGGTACCCCAAG 2594  
Db 2944 CTTTATTAAGAAATTCCTTAGCGTGTGTTTCTCGAAATGCTCGCGGTACCCCAAG 3003  
Qy 2595 ATTAGAAATGAGTAGATCAAAATTTTACCAATAGAAATCAGGAAATTCAGATCCAAACCAT 2654  
Db 3004 ATTAGAAATGAGTAGATCAAAATTTTACCAATAGAAATCAGGAAATTCAGATCCAAACCAT 3063  
Qy 2655 AAAAACHACTAGACAAATTTGCAAGTTAACTACTCAAGCTAGTAGTGGATTTATCCC 2714  
Db 3064 AAAAACHACTAGACAAATTTGCAAGTTAACTACTCAAGCTAGTAGTGGATTTATCCC 3123  
Qy 2715 AAATGAGCCACAGAACCCAGAGCCAGAAACAGAAATCAGAACTAACTATGAGTTTGA 2774  
Db 3124 AAATGAGCCACAGAACCCAGAGCCAGAAACAGAAATCAGAACTAACTATGAGTTTGA 3183

Qy 2775 AATGGAAGAGAAAAAGCAATGACTTCTGTGAATAATGACGAAATCGTTGCTTATTT 2834  
Db 3184 AATGGAAGAGAAAAAGCAATGACTTCTGTGAATAATGACGAAATCGTTGCTTATTT 3243  
Qy 2835 TTTTATAAAGCGGTATACGTATATAACGAAACAAACGAACTGATAGAAACGAAAG 2894  
Db 3244 TTTTATAAAGCGGTATACGTATATAACGAAACAAACGAACTGATAGAAACGAAAG 3303  
Qy 2895 AGCCATGACACATTTATAAAATGTTTGAACGACATTTTATAAATCATAGCCCGATAAGAT 2954  
Db 3304 AGCCATGACACATTTATAAAATGTTTGAACGACATTTTATAAATCATAGCCCGATAAGAT 3363  
Qy 2955 TGCMAAACCAACGCTTATACGTATAGTACGATGAATCTCTTCCCTCGTAAAGATTTATTAA 3014  
Db 3364 TGCMAAACCAACGCTTATACGTATAGTACGATGAATCTCTTCCCTCGTAAAGATTTATTAA 3423  
Qy 3015 TTAACCTTTGTTTGAAGCGGTATATAACCGTACTATCATTTATATAGGGAATCAGAGAT 3074  
Db 3424 TTAACCTTTGTTTGAAGCGGTATATAACCGTACTATCATTTATATAGGGAATCAGAGAT 3483  
Qy 3075 TTTCAAGTATCTAAGCTACTCAATTTAAGAAATGTTAAGCAATCAATCGGAAATCGTTT 3134  
Db 3484 TTTCAAGTATCTAAGCTACTCAATTTAAGAAATGTTAAGCAATCAATCGGAAATCGTTT 3543  
Qy 3135 ATTGCTTTTGTATTTTATTTATAGAGGTGGAGTTTGTATGATCATGATGATGATGAA 3194  
Db 3544 ATTGCTTTTGTATTTTATTTATAGAGGTGGAGTTTGTATGATCATGATGATGATGAA 3603  
Qy 3195 AACTTATATAAAAAATAGTTTATTTGGAGATAAGAAAAATTAGCAATATCTATACACTAGA 3254  
Db 3604 AACTTATATAAAAAATAGTTTATTTGGAGATAAGAAAAATTAGCAATATCTATACACTAGA 3663  
Qy 3255 AACGTTAAGAAAGAGTTAGAAAAAGAGAAATATCTACTAGAAAAAATAATCAGATAGTA 3314  
Db 3664 AACGTTAAGAAAGAGTTAGAAAAAGAGAAATATCTACTAGAAAAAATAATCAGATAGTA 3723  
Qy 3315 TTTTCTTTCGGAGGGGAGATTTATATATAAGTTTAAAGAAATAACAAAAATAATTTA 3374  
Db 3724 TTTTCTTTCGGAGGGGAGATTTATATATAAGTTTAAAGAAATAACAAAAATAATTTA 3783  
Qy 3375 TTCGATTAGTGGAAAAAATTTGACTTATAAGGAAAAAATCTTTTCAAAAACATGCAAT 3434  
Db 3784 TTCGATTAGTGGAAAAAATTTGACTTATAAGGAAAAAATCTTTTCAAAAACATGCAAT 3843  
Qy 3435 ATTGAAACAGTTGATGAAAGCAACCAAGTTTAAATTAACAAACCTATTTTATAGGAT 3494  
Db 3844 ATTGAAACAGTTGATGAAAGCAACCAAGTTTAAATTAACAAACCTATTTTATAGGAT 3903  
Qy 3495 TATAGAAAGGAGAACAGCTGAATGAATATCCCTTTTGTGTAGAAATCTGCTTCAATGA 3554  
Db 3904 TATAGAAAGGAGAACAGCTGAATGAATATCCCTTTTGTGTAGAAATCTGCTTCAATGA 3963  
Qy 3555 CGGCTTGTAAAGTACAAATTTAAATAATGTAATAATTCGCTCAATCTACTACCAAGCCAGG 3614  
Db 3964 CGGCTTGTAAAGTACAAATTTAAATAATGTAATAATTCGCTCAATCTACTACCAAGCCAGG 4023  
Qy 3615 TAAACAGAAAGGGGCTATTTTGTGTATCGCTCAAAATCAAGCATGATTTGCGGTGCTGG 3674  
Db 4024 TAAACAGAAAGGGGCTATTTTGTGTATCGCTCAAAATCAAGCATGATTTGCGGTGCTGG 4083  
Qy 3675 TGTGTTCTGACTTCCGAGGAGCGATTCAGAGAAATCAAGATACATTTACATTTGAC 3734  
Db 4084 TGTGTTCTGACTTCCGAGGAGCGATTCAGAGAAATCAAGATACATTTACATTTGAC 4143  
Qy 3735 ACCCAACGTTTATCGTTATGGAACGTATGAGAGAAACCGTTTATACACAGAAAGGACA 3794  
Db 4144 ACCCAACGTTTATCGTTATGGAACGTATGAGAGAAACCGTTTATACACAGAAAGGACA 4203  
Qy 3795 TTTCTGAAACCAATTTAAGACAAATCAATTCCTTTTATTTGATTTTATTTTACACGGC 3854  
Db 4204 TTTCTGAAACCAATTTAAGACAAATCAATTCCTTTTATTTGATTTTATTTTACACGGC 4263



QY 559 TGAGCAATTAAGTAAACCCCTTGGGGCTCTAAACGGGCTTGGGGGTTTTTGTCT 618  
DB 932 TGAGCAATTAAGTAAACCCCTTGGGGCTCTAAACGGGCTTGGGGGTTTTTGTCT 991  
QY 619 GAAAGAGGAAGTATATCCGGATGACCTGAGGCAAGCTAGAAATCGATACGATTTTGA 678  
DB 992 GAAAGAGGAAGTATATCCGGATGACCTGAGGCAAGCTAGAAATCGATACGATTTTGA 1051  
QY 679 AGTGGCAACAGATAAATAAAGCAGTTTAAATTTGTGTGAACTTTTAAACAAGCAAA 738  
DB 1052 AGTGGCAACAGATAAATAAAGCAGTTTAAATTTGTGTGAACTTTTAAACAAGCAAA 1111  
QY 739 TACAATCATTTGTCGCAACAGATAGCGACAGAGAGCGGAAACATTTGCTGCGATCAT 798  
DB 1112 TACAATCATTTGTCGCAACAGATAGCGACAGAGAGCGGAAACATTTGCTGCGATCAT 1171  
QY 799 TCATAAAGCAAAATGCCCTTTCTAAAGATAAAGCGTATAAAGAGCTATGGATCAATAGTTT 858  
DB 1172 TCATAAAGCAAAATGCCCTTTCTAAAGATAAAGCGTATAAAGAGCTATGGATCAATAGTTT 1231  
QY 859 AGAAAGAGTGTATCCGTAGCGGTTTCAAAATTTGCAACAGGAATGAATTAATCTATCC 918  
DB 1232 AGAAAGAGTGTATCCGTAGCGGTTTCAAAATTTGCAACAGGAATGAATTAATCTATCC 1291  
QY 919 CTTTTATCAAGAGCGCAAAAGAAACGAAATGATACCAATCAGTGCAGAAAGAAAT 978  
DB 1292 CTTTTATCAAGAGCGCAAAAGAAACGAAATGATACCAATCAGTGCAGAAAGAAAT 1351  
QY 979 ATAATGGAGATAGAGCGGTTTGGTTCGTGCTGATGTCACCATATCAATAAATAATCGAA 1038  
DB 1352 ATAATGGAGATAGAGCGGTTTGGTTCGTGCTGATGTCACCATATCAATAAATAATCGAA 1411  
QY 1039 ACAGCAAAAGATGGCGAAACGTTAAAGAGTTATGGAATAGACTTAGAGCAAACTT 1098  
DB 1412 ACAGCAAAAGATGGCGAAACGTTAAAGAGTTATGGAATAGACTTAGAGCAAACTT 1471  
QY 1099 AAGAGTGTGTAGTATGTCAGTATCTTAAATTTGTATAATAGGAATGAAGTTAAAT 1158  
DB 1472 AAGAGTGTGTAGTATGTCAGTATCTTAAATTTGTATAATAGGAATGAAGTTAAAT 1531  
QY 1159 AGATGCTAAAATTTGTATAGAGGAGTGTATACATGACCAAAATATATAATATTC 1218  
DB 1532 AGATGCTAAAATTTGTATAGAGGAGTGTATACATGACCAAAATATATAATATTC 1591  
QY 1219 TCAAACTTTTAAACGAGTGAAAGAGTACTCAACCAAAATAAATAAACAATTTGAAATTTAA 1278  
DB 1592 TCAAACTTTTAAACGAGTGAAAGAGTACTCAACCAAAATAAATAAACAATTTGAAATTTAA 1651  
QY 1279 AGAAACGATACCGTTTACGAATTTGGACAGGTAAAGGCAATTTAACGACGAACTGGC 1338  
DB 1652 AGAAACGATACCGTTTACGAATTTGGACAGGTAAAGGCAATTTAACGACGAACTGGC 1711  
QY 1339 TAAATTAAGTAAACAGGTAAAGTCTATTTGAATTTAGACAGTCTATTTCAACTTATCGTC 1398  
DB 1712 TAAATTAAGTAAACAGGTAAAGTCTATTTGAATTTAGACAGTCTATTTCAACTTATCGTC 1771  
QY 1399 AGAAAAATTAATCTGATCTCGTGTCACTTTTAAATTTCAACCAAGATTTCTACAGTTTCA 1458  
DB 1772 AGAAAAATTAATCTGATCTCGTGTCACTTTTAAATTTCAACCAAGATTTCTACAGTTTCA 1831  
QY 1459 ATTCCCTTAAACAGAGGTATTAATTTGTGGAGTATTTCCATTACCAATTTAGGACACA 1518  
DB 1832 ATTCCCTTAAACAGAGGTATTAATTTGTGGAGTATTTCCATTACCAATTTAGGACACA 1891  
QY 1519 AATTATTAAAAAGTGGTTTTTGAAGCCATGCGTCTGACATCTATCTGATTTGTGAAGA 1578  
DB 1892 AATTATTAAAAAGTGGTTTTTGAAGCCATGCGTCTGACATCTATCTGATTTGTGAAGA 1951  
QY 1579 AGAATTTCAAGGTTACCTTGAATTTACCGAATTTACCGAATTTACCGAATTTACCGAAT 1638  
DB 1952 AGAATTTCAAGGTTACCTTGAATTTACCGAATTTACCGAATTTACCGAATTTACCGAAT 2011

QY 1639 AGTCTCGATTGCAAAATGCTTAAAGTCCAGCGGAATGCTTTTATCTTAAACCAAGT 1698  
DB 2012 AGTCTCGATTGCAAAATGCTTAAAGTCCAGCGGAATGCTTTTATCTTAAACCAAGT 2071  
QY 1699 AAAACAGTCTCTTAATAAATTTACCGCATACACAGATGTTCCAGATAAATATTGAA 1758  
DB 2072 AAAACAGTCTCTTAATAAATTTACCGCATACACAGATGTTCCAGATAAATATTGAA 2131  
QY 1759 GCTATATACGTACTTGTGTTTCAAAATGCGTCAATTCGAGATATCTGCTCACTGTTACTAA 1818  
DB 2132 GCTATATACGTACTTGTGTTTCAAAATGCGTCAATTCGAGATATCTGCTCACTGTTACTAA 2191  
QY 1819 AAATCAGTCTTATCAAGCAATGAAACACGCGCAAAAGTAAACAAATTTAAGTACCGTTACTTA 1878  
DB 2192 AAATCAGTCTTATCAAGCAATGAAACACGCGCAAAAGTAAACAAATTTAAGTACCGTTACTTA 2251  
QY 1879 TGAGCAAGTATGCTCTATTTTAAATGATTTATTTAAACGGGAGGAATTAATTTCTA 1938  
DB 2252 TGAGCAAGTATGCTCTATTTTAAATGATTTATTTAAACGGGAGGAATTAATTTCTA 2311  
QY 1939 TGAGTCCGCTTTTGTAAATTTGAAAGTTTACAGTTTACTTAAAGGGAATGTAGATAAATTTAT 1998  
DB 2312 TGAGTCCGCTTTTGTAAATTTGAAAGTTTACAGTTTACTTAAAGGGAATGTAGATAAATTTAT 2371  
QY 1999 TAGGTATATCTACTGACAGCTTTCCAAAGAGCTTAAAGAGTCCCTAGCGCTCTTATCATGG 2058  
DB 2372 TAGGTATATCTACTGACAGCTTTCCAAAGAGCTTAAAGAGTCCCTAGCGCTCTTATCATGG 2431  
QY 2059 GAAGCTCGGATCATATGCAAGCAAAATTAACCTGCAACAGCACTTGGAGAAATGGGACG 2118  
DB 2432 GAAGCTCGGATCATATGCAAGCAAAATTAACCTGCAACAGCACTTGGAGAAATGGGACG 2491  
QY 2119 AATCGAGAAAAACCTCTTTACGCTGGATTACATATCTAATAAAGCCGTAAAGAGACGGGT 2178  
DB 2492 AATCGAGAAAAACCTCTTTACGCTGGATTACATATCTAATAAAGCCGTAAAGAGACGGGT 2551  
QY 2179 TCAAAAGGTTTAAATAAAGAGAGCAATCAATGCAATTAGCTAGACTATATTTTTTGG 2238  
DB 2552 TCAAAAGGTTTAAATAAAGAGAGCAATCAATGCAATTAGCTAGACTATATTTTTTGG 2611  
QY 2239 ACAACGTCGAGAAATTTAGAGAACGTCCTCTCCAAGACAGTTTACAAAGAGCTAGTGCAC 2298  
DB 2612 ACAACGTCGAGAAATTTAGAGAACGTCCTCTCCAAGACAGTTTACAAAGAGCTAGTGCAC 2671  
QY 2299 AAAACATAATTTAAACGCTATTAAGTGTGTGAAACACTGTATATATGGAAGAAAGCGTAGA 2358  
DB 2672 AAAACATAATTTAAACGCTATTAAGTGTGTGAAACACTGTATATATGGAAGAAAGCGTAGA 2731  
QY 2359 AGAATTAAGCAACAGAGAGAAATTTAGAGAGATTTAATGCCATATCGTGGCGGTTAGG 2418  
DB 2732 AGAATTAAGCAACAGAGAGAAATTTAGAGAGATTTAATGCCATATCGTGGCGGTTAGG 2791  
QY 2419 ATGGGAACATATCAATTTTCTTGAGGAATAAATAATTTGAAGGATTCATGACACTGGGCA 2478  
DB 2792 ATGGGAACATATCAATTTTCTTGAGGAATAAATAATTTGAAGGATTCATGACACTGGGCA 2851  
QY 2479 AATGAATTTACGCTCTTACGATTAAGAGCCGTTTTTCTTAATATACCGCTCTTT 2538  
DB 2852 AATGAATTTACGCTCTTACGATTAAGAGCCGTTTTTCTTAATATACCGCTCTTT 2911  
QY 2539 TTATAGAAAAAATCCTTACGCTGTTTTTTTCCGAAATGCTGGCGGTACCCCAAGAAATTA 2598  
DB 2912 TTATAGAAAAAATCCTTACGCTGTTTTTTTCCGAAATGCTGGCGGTACCCCAAGAAATTA 2971  
QY 2599 GAAATGAGTAGATCAAAATTTCTCGAATAGAAATCAGGAAATCAGATCCCAACATAAAA 2658  
DB 2972 GAAATGAGTAGATCAAAATTTCTCGAATAGAAATCAGGAAATCAGATCCCAACATAAAA 3031  
QY 2659 ACCTTAGAACAATTTGCAAGTAACTAACTCAACGCTAGTAGTGGATTTAATCCCAAT 2718  
DB 3032 ACCTTAGAACAATTTGCAAGTAACTAACTCAACGCTAGTAGTGGATTTAATCCCAAT 3091  
QY 2719 GAGCCAAACAGAACCCAGAGCCAGAAAAACAGAAATCAGAACAGTAACTTGGATTTTGAATG 2778







Db 1308 AAGAGGATTCTACAGCGTACCTTGGATATTACCCGNACTAGGGTTGCTCTTGCA 1367  
Qy 1635 CTCAGTCTCGATTACAGCAATTGCTTAAGCTGCCAGGGGAATGCTTTCATCTCAACCAA 1694  
Db 1368 CTCAGTCTCGATTACAGCAATTGCTTAAGCTGCCAGGGGAATGCTTTCATCTCAACCAA 1427  
Qy 1695 AAGTAAACAGTGTCTTAATAAACTTACCGCCATACACAGATGTTTCCAGATATAAT 1754  
Db 1428 AAGTAAACAGTGTCTTAATAAACTTACCGCCATACACAGATGTTTCCAGATATAAT 1487  
Qy 1755 GGAAGCTATATACGTAATTTGTTTCAAAATGGGTCAATCGAGANAATCGTCAACTGTTTA 1814  
Db 1488 GGAAGCTATATACGTAATTTGTTTCAAAATGGGTCAATCGAGANAATCGTCAACTGTTTA 1547  
Qy 1815 CTAATAATCAGTTTCATCAAGCAATGAACACGCAAGTAATAACAAATTTAAGTACCGTTA 1874  
Db 1548 CTAATAATCAGTTTCATCAAGCAATGAACACGCAAGTAATAACAAATTTAAGTACCGTTA 1607  
Qy 1875 CTTATGACCAAGTATGTTCTATTTTAAATAGTTATCTATTATTTAAACGGGAGGAATAAT 1934  
Db 1608 CTTATGACCAAGTATGTTCTATTTTAAATAGTTATCTATTATTTAAACGGGAGGAATAAT 1667  
Qy 1935 TCTATGAGTCGTTTGTAAATTTGGAAGTTACAGTTACTTAAGGGAATGTAGATAAA 1994  
Db 1668 TCTATGAGTCGTTTGTAAATTTGGAAGTTACAGTTACTTAAGGGAATGTAGATAAA 1727  
Qy 1995 TTATTAGGTATACTACTGACAGCTTCCAGAGGCTTAAAGAGTCCCTAGCGCTCTTATCA 2054  
Db 1728 TTATTAGGTATACTACTGACAGCTTCCAGAGGCTTAAAGAGTCCCTAGCGCTCTTATCA 1787  
Qy 2055 TGGGAAGCTCGATCATATCAAGACAAATAAATCTGCAACAGCACTTGGGAATATG 2114  
Db 1788 TGGGAAGCTCGATCATATCAAGACAAATAAATCTGCAACAGCACTTGGGAATATG 1847  
Qy 2115 CACGAATCGGAACAACTCTTTACGTTGGATTACATATCTAATAAGCCGTAAAGAGAC 2174  
Db 1848 CACGAATCGGAACAACTCTTTACGTTGGATTACATATCTAATAAGCCGTAAAGAGAC 1907  
Qy 2175 GGGTTCAAAAAGTTTAAATAAGGAGAGCAATCAATGCAATTAGCTAGAACATATATTT 2234  
Db 1908 GGGTTCAAAAAGTTTAAATAAGGAGAGCAATCAATGCAATTAGCTAGAACATATATTT 1967  
Qy 2235 TTGACACAGTGGAGATTTAGAGACGTCCTCTCCAGCCGTTACRAAGAGCTAGTG 2294  
Db 1968 TTGACACAGTGGAGATTTAGAGACGTCCTCTCCAGCCGTTACRAAGAGCTAGTG 2027  
Qy 2295 CACTAAACATATTAATTAACGCTATAAGTGTGTGGAACACCTGTATATATGGAAGAGCCG 2354  
Db 2028 CACTAAACATATTAATTAACGCTATAAGTGTGTGGAACACCTGTATATATGGAAGAGCCG 2087  
Qy 2355 TAGAAGAAATTAAGCAAGAGGAGAAATTAGAGAGATTTAAATGCCATATGCGTGGCCGT 2414  
Db 2088 TAGAAGAAATTAAGCAAGAGGAGAAATTTAGAGAGATTTAAATGCCATATGCGTGGCCGT 2147  
Qy 2415 TAGATGGGAACATATCAATTTCTCGAGAAATCAAAATTTGAAGGATTAACATGACACTG 2474  
Db 2148 TAGATGGGAACATATCAATTTCTCGAGAAATCAAAATTTGAAGGATTAACATGACACTG 2207  
Qy 2475 GCGAAATGAATTAACGTCCTTACGTATAAAGAGCGGTTTATTTCTTAATATAACCGCT 2534  
Db 2208 GCGAAATGAATTAACGTCCTTACGTATAAAGAGCGGTTTATTTCTTAATATAACCGCT 2267  
Qy 2535 CTTTTTATAGAAAAAATCCTTTAGCGTGGTTTTTTTCCGAAATGCTGCGGTACCCCAAGA 2594  
Db 2268 CTTTTTATAGAAAAAATCCTTTAGCGTGGTTTTTTTCCGAAATGCTGCGGTACCCCAAGA 2327  
Qy 2595 ATTAGAAATGATAGATCAAAATTTACGAAATAGAAATCAGGAAATCAGATCAACCAT 2654  
Db 2328 ATTAGAAATGATAGATCAAAATTTACGAAATAGAAATCAGGAAATCAGATCAACCAT 2387  
Qy 2655 AABAACACTAGAACAAATTCGAAAGTTAACTTAACCTCAACCGCTAGTAGGATTTAATCCC 2714  
Db 2388 AABAACACTAGAACAAATTCGAAAGTTAACTTAACCTCAACCGCTAGTAGGATTTAATCCC 2447

Qy 2715 AAATGAGCCACAGAACAGAGCCAGAACAGNAATCAGAACAGCTTAACATTTGGATTAGA 2774  
Db 2448 AAATGAGCCACAGAACAGAGCCAGAACAGNAATCAGAACAGCTTAACATTTGGATTAGA 2507  
Qy 2775 AATGGAAGAAAGAAAAAGCAATGACTTCGTGTGAATAATGACGAAATCGTGTCTATTT 2834  
Db 2508 AATGGAAGAAAGAAAAAGCAATGACTTCGTGTGAATAATGACGAAATCGTGTCTATTT 2567  
Qy 2835 TTTTAAAAAGCGGTATACTAGATATAACGAAACACGAACTGAATAGAACGAAAAAG 2894  
Db 2568 TTTTAAAAAGCGGTATACTAGATATAACGAAACACGAACTGAATAGAACGAAAAAG 2627  
Qy 2895 AGCCATGACACATTTATAAAATGTTTGAAGCACTTTATAAATGCATAGCCCGGATAGAT 2954  
Db 2628 AGCCATGACACATTTATAAAATGTTTGAAGCACTTTATAAATGCATAGCCCGGATAGAT 2687  
Qy 2955 TGCCAAAACCAAGCTTATCAGTTAGTCAGATGAACCTTCCCTCGTAAAGAGTTATTTAA 3014  
Db 2688 TGCCAAAACCAAGCTTATCAGTTAGTCAGATGAACCTTCCCTCGTAAAGAGTTATTTAA 2747  
Qy 3015 TTAACTTTTGTGTAAGACGCTATATAACCGTACTATCATATATATAGGGAATCAGAGAT 3074  
Db 2748 TTAACTTTTGTGTAAGACGCTATATAACCGTACTATCATATATATAGGGAATCAGAGAT 2807  
Qy 3075 TTTCAGTATCTAAGCTACTGAATTTAAAGAAATGTTTAAAGCAATCAATCGGAAATCGTTTG 3134  
Db 2808 TTTCAGTATCTAAGCTACTGAATTTAAAGAAATGTTTAAAGCAATCAATCGGAAATCGTTTG 2867  
Qy 3135 ATTGCTTTTGTGTAATTCATTTATAGAGGTGGAGTTTGTATGAATCATGATGAATGTAA 3194  
Db 2868 ATTGCTTTTGTGTAATTCATTTATAGAGGTGGAGTTTGTATGAATCATGATGAATGTAA 2927  
Qy 3195 AACTTTATATAAAAAATAGTTTATTGGAGATAAGAAAAATTAGCAAAATATCTATACACTAGA 3254  
Db 2928 AACTTTATATAAAAAATAGTTTATTGGAGATAAGAAAAATTAGCAAAATATCTATACACTAGA 2987  
Qy 3255 AACGTTTAAAGAAAGTTAGAAAAAGAGAAATATCTCTTAGAAAAAATAACAGATAGTA 3314  
Db 2988 AACGTTTAAAGAAAGTTAGAAAAAGAGAAATATCTCTTAGAAAAAATAACAGATAGTA 3047  
Qy 3315 TTTTCTTTCGAGGGGGAGATTTATATATAAGTTTAAATAGAAAAATAACAAAAATTTTA 3374  
Db 3048 TTTTCTTTCGAGGGGGAGATTTATATATAAGTTTAAATAGAAAAATAACAAAAATTTTA 3107  
Qy 3375 TTCGATTAGTGGAAAAAATTTGACTTTATAAGGAAAAAATCTTTTTCAAAAACATGCAAT 3434  
Db 3108 TTCGATTAGTGGAAAAAATTTGACTTTATAAGGAAAAAATCTTTTTCAAAAACATGCAAT 3167  
Qy 3435 ATTGAAACAGTTGAATGAAAAAGCAACCAAGTTTAAATTAACCAACCTTTTATAGGAT 3494  
Db 3168 ATTGAAACAGTTGAATGAAAAAGCAACCAAGTTTAAATTAACCAACCTTTTATAGGAT 3227  
Qy 3495 TATAGGAAGGAGAACAGCTGAATATATCCCTTTTGTGTAGAAACTGTCCTTCATGA 3554  
Db 3228 TATAGGAAGGAGAACAGCTGAATATATCCCTTTTGTGTAGAAACTGTCCTTCATGA 3287  
Qy 3555 CGGCTTGTAAAGTCAAAATTTAAAAATAGTAAATTCGCTCAATCACTACCAAGCCAGG 3614  
Db 3288 CGGCTTGTAAAGTCAAAATTTAAAAATAGTAAATTCGCTCAATCACTACCAAGCCAGG 3347  
Qy 3615 TAAAGCAAAAGGGGCTATTTTTCGCTATCGCTCAAAATCAAGCATGATTTGCGGTCTGG 3674  
Db 3348 TAAAGCAAAAGGGGCTATTTTTCGCTATCGCTCAAAATCAAGCATGATTTGCGGTCTGG 3407  
Qy 3675 TGTTGTTCTGACTCCGAGGAGAGGATTCAGAAAAATCAAGATACATTTACACATTCGAC 3734  
Db 3408 TGTTGTTCTGACTCCGAGGAGGATTCAGAAAAATCAAGATACATTTACACATTCGAC 3467  
Qy 3735 ACCCAACGTTTATCGTTATGGAAGTTATGCAAGCAAAACCGTTTCATACACGAAAGACA 3794  
Db 3468 ACCCAACGTTTATCGTTATGGAAGTTATGCAAGCAAAACCGTTTCATACACGAAAGACA 3527



4608	db	AGAAAGCTATATAAAGGCGCTGACAAATTCCTTTGACTTTAGAGCATACATTCTTTCRAAGA	4667
4935	QY	GACTTTAAACAAGCTAGCAGAACGCCCTAAACACGGACACACAACCTCGATTTGCTTTTAGCTA	4994
4668	Db	GACTTTAAACAAGCTAGCAGAACGCCCTAAACACGGACACACAACCTCGATTTGCTTTAGCTA	4727
4995	QY	TGATACAGGCTGAAAATAAAACCCGCACTATGCCATTACATTATATCTATGATACGCTGT	5054
4728	Db	TGATACAGGCTGAAAATAAAACCCGCACTATGCCATTACATTATATCTATGATACGCTGT	4787
5055	QY	TTGTGTTTTCTTTGCTGTTTTAGCGAATGATTAGCAGAAATATACAGAGTAAGATTTTAAT	5114
4788	Db	TTGTGTTTTCTTTGCTGTTTTAGCGAATGATTAGCAGAAATATACAGAGTAAGATTTTAAT	4847
5115	QY	TAATTTATTAGGGGGAGAGGAGAGTAGCCCGAAAACCTTTTAGTTGGCTTGGACTGAC	5174
4848	Db	TAATTTATTAGGGGGAGAGGAGAGTAGCCCGAAAACCTTTTAGTTGGCTTGGACTGAC	4907
5175	QY	GAAGTGAGGGAAAGGCTACTAAAACGTCGAGGGCAGTGAGAGCGAAAGCAACCTTGAT	5234
4908	Db	GAAGTGAGGGAAAGGCTACTAAAACGTCGAGGGCAGTGAGAGCGAAAGCAACCTTGAT	4967
5235	QY	TTTTTAAATTTCTATCTTTTAGTCAITAGAGTATACTTATTTGTGCTATAAACAATT	5294
4968	Db	TTTTTAAATTTCTATCTTTTAGTCAITAGAGTATACTTATTTGTGCTATAAACAATT	5027
5295	QY	TAGCAGCATAAATAGATTTATCGAATAGGTCATTTAAGTTTGACATATTAGAGGAGAAAA	5354
5028	Db	TAGCAGCATAAATAGATTTATCGAATAGGTCATTTAAGTTTGACATATTAGAGGAGAAAA	5087
5355	QY	TCTTGAGAAATATTTTGAAGAACCCGATTACATGATTGGATTAGTTCTTTGGTTACGT	5414
5088	Db	TCTTGAGAAATATTTTGAAGAACCCGATTACATGATTGGATTAGTTCTTTGGTTACGT	5147
5415	QY	GGTTTTTAACATAAAGTAGTGAATTTTGAATTTTGGTGTGTGTCTTGTTGTTTAGTAT	5474
5148	Db	GGTTTTTAACATAAAGTAGTGAATTTTGAATTTTGGTGTGTGTCTTGTTGTTTAGTAT	5207
5475	QY	TTGCTAGTCAAAAGTGATTTAAATA	5497
5208	Db	TTGCTAGTCAAAAGTGATTTAAATA	5230

RESULT 5  
US-09-070-927A-287/c  
; Sequence 287, Application US/09070927A  
; Patent No. US20020120116A1  
; GENERAL INFORMATION:  
; APPLICANT: Charles A. Kunsch  
; Patrick J. Dillion  
; Steven Barash  
; TITLE OF INVENTION: Enterococcus faecialis polynucleotides and polypeptides

```

US-09-070-927A-287/c
; Sequence 287, Application US/09070927A
; Patent No. US20020120116A1
; GENERAL INFORMATION:
; APPLICANT: Charles A. Kunsch
; Patrick J. Dillon
; Steven Barash
; TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 982
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/070,927A
; FILING DATE: 04-May-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/046,655
; FILING DATE: 1997-05-16

```

APPLICATION NUMBER: 60/044,031  
 FILING DATE: 1997-05-06  
 APPLICATION NUMBER: 60/066,009  
 FILING DATE: 1997-11-14  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Kenley K. Hoover  
 REGISTRATION NUMBER: 40,302  
 REFERENCE/DOCKET NUMBER: PB369  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (301) 309-8504  
 TELEFAX: (301) 309-8512  
 INFORMATION FOR SEQ ID NO: 287:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 3765 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 SEQUENCE DESCRIPTION: SEQ ID NO: 287:  
 US-09-070-927A-287

Query Match 26.1%; Score 1434.8; DB 9; Length 3765;

Best local Similarity 96.6%; Pred. No. 3.4e-270; Indels 2; Gaps 2;  
 Matches 1485; Conservative 2; Mismatches 49;

QY	3488	TAGGATTTATAGGAAGGAGAACAGCTCAATGAATATCCCTTTTGGTTGTAGAACTGTGC	3547
DB	2075	TAATTTTCGTAAGAAGGAGAACAGCTCAATGAATATCCCTTTTGGTTGTAGAACTGTGC	2016
QY	3548	TTATGACGGCTGTGTTAAAGTACAAATTTAAATAATAGTAAATTCGCTCAATCACTACCA	3607
DB	2015	TTATGACGGCTGTGTTAAAGTACAAATTTAAATAATAGTAAATTCGCTCAATCACTACCA	1956
QY	3608	AGCCAGGTAAAGCAAGGAGGCTATTTTGGTATCGCTCAAAATCAAGCATGATCGCG	3667
DB	1955	AGCCAGGTAAAGCAAGGAGGCTATTTTGGTATCGCTCAAAATCAAGCATGATCGCG	1896
QY	3668	GTCGTGGTGTGTTCTGACTTCCGAGGAGCGATTCAGAAATCAAGATACATTTACAC	3727
DB	1895	GACGTGGGCTGTGTTCTGACTTCCGAGGAGCGATTCAGAAATCAAGATACATTTACGC	1836
QY	3728	ATTGACACCCAAAGCTTTATCGTTATGGAAGTATGACGAGCAAAACCGTTTCATACGA	3787
DB	1835	ATTGACACCCAAAGCTTTATCGTTATGGAAGTATGACGAGCAAAACCGTTTCATACGA	1776
QY	3788	AAGGACATCTGAAACAAATTTAAGACAAATCAATACCTTTCTTTATTTGATTTGATATTC	3847
DB	1775	AAGGACATCTGAAACAAATTTAAGACAAATCAATACCTTTCTTTATTTGATTTGATATTC	1716
QY	3848	ACAGGCAAAAGAACTATTTTCAGCAAGCGATTTTAAACACCGCTATTGATTAGGTT	3907
DB	1715	ACAGGCAAAAGAACTATTTTCAGCAAGCGATTTTAAACACCGCTATTGATTAGGTT	1656
QY	3908	TTATGCCCTACTGATGATATCAATCTGATAAGGTTATCAAGCATATTTTGGTTTAGAAA	3967
DB	1655	TTATGCCCTACTGATGATATCAATCTGATAAGGTTATCAAGCATATTTTGGTTTAGAAA	1596
QY	3968	CGCAGTCTATGTGACTTCAAAATCAGATTTAAATCTGTCAAGAGCGCAAAATATTTT	4027
DB	1595	CGCCAGT-TATGTGACTTCAAAATCAGATTTAAATCTGTCAAGAGCGC-AAAATATCT	1538
QY	4028	CGCAAAATATCCGAGATATTTTGGAAAGCTTTTGGCCAGTTGATCTAACGTGTAATCAT	4087
DB	1537	CGCAAAATATCCGAGATATTTTGGAAAGCTTTTGGCCAGTTGATCTAACGTGTAATCAT	1478
QY	4088	TTGATTTGCTCGCATACCAAGACGCAATATGTAGAAATTTTGTATTCCTTAATTCGGTT	4147
DB	1477	TTGGGAAWGTCTGATATACCAAGACGCAATATGTAGAAATTTTGTATTCCTTAATTCGGTT	1418
QY	4148	ATTCTTTCAAGAAATGGCAAGATTTGCTTTTCAACAAACAGATATAAGGCTTTTACTC	4207
DB	1417	ATTCTTTCAAGAAATGGCAAGATTTGCTTTTCAACAAACAGATATAAGGCTTTTACTC	1358
QY	4208	GTTCAAGTCTAACGGTTTAAAGCGGTACAGAGGCAAAACCAAGTATGATGAACCCCTGGT	4267

DB	1357	GTTCAAGTCTAACGGTTTAAAGCGGTACAGAGGCAAAACCAAGTAGTGAACCCCTGGT	1298
QY	4268	TTAATCTCTTATTCACGAAACGAAATTTTCAGGAGAAAGGGTTTAAATAGGGCGTATA	4327
DB	1297	TTAATCTCTTATTCACGAAACGAAATTTTCAGGAGAAAGGGTTTAAATAGGGCGTATA	1238
QY	4328	ACGTCAATGTTACCTCTCTTTAGCCTACTTTAGTTCAGGCTATTCATTCGAAACGTCG	4387
DB	1237	ACGTCAATGTTACCTCTCTTTAGCCTACTTTAGTTCAGGCTATTCATTCGAAACGTCG	1178
QY	4388	AATATAATATGTTTGAGTTTAAATATCAGTAGATCAACCCCTTAGAAGAAAGAAAGTAA	4447
DB	1177	AATATAATATGTTTGAGTTTAAATATCAGTAGATCAACCCCTTAGAAGAAAGAAAGTAA	1118
QY	4448	TCAAAATGTTTAGAAGTGCCTTATTCAGAAACTATCAAGGGGCTTAATAGGGAATACATTA	4507
DB	1117	TTAAACTTGTAGAAAGTGCCTTATTCAGAAACTATCAAGGGGCTTAATAGGGAATACATTA	1058
QY	4508	CCATCTCTTGCRAAGCTTGGGTATCAAGTATTAACAGTAAAGATTTTGTCCGTC	4567
DB	1057	CCATCTCTTGCRAAGCTTGGGTATCAAGTATTAACAGTAAAGATTTTGTCCGTC	998
QY	4568	AAGGCTGGTTTAAATTCAGAAAGAAAGAGCGAAGTCAACGTTTCATTTGTCAGAAAT	4627
DB	997	AAGGCTGGTTTAAATTCAGAAAGAAAGAGCGAAGTCAACGTTTCATTTGTCAGAAAT	938
QY	4628	CGAAAGAGATTTAATGCTTATATTCGAAAGAAAGCGATGTATACAGGCTTTATTAG	4687
DB	937	CGAAAGAGATTTAATGCTTATATTCGAAAGAAAGCGATGTATACAGGCTTTATTAG	878
QY	4688	TCACCAACCAAAAGAGATTTAGAGAAAGTGTAGGCAATTCCTGAAACGGACATTAGATAAT	4747
DB	877	TCACCAACCAAAAGAGATTTAGAGAAAGTGTAGGCAATTCCTGAAACGGACATTAGATAAT	818
QY	4748	TCCTGAAAGTACTGAAGGCGAATCAGGAAATTTCTTTAAGATTTAAACCGAGAGAAATG	4807
DB	817	TCCTGAAAGTACTGAAGGCGAATCAGGAAATTTCTTTAAGATTTAAACCGAGAGAAATG	758
QY	4808	GTGGCAATCAACTCTGCTAGTGTAAATCATTTGCTATCGATCATTTAAAGTAAAAAAG	4867
DB	757	GTGGCAATCAACTCTGCTAGTGTAAATCATTTGCTATCGATCATTTAAAGTAAAAAAG	698
QY	4868	AAGAAAAAGAAAGCTATATAAGGCGCTGACAAATTTCTTTGACTTAGAGCATACATTC	4927
DB	697	AAGAAAAAGAAAGCTATATAAGGCGCTGTCAGAGTTTTTTGACTTAGAGCATACATTC	638
QY	4928	TTCAAGAGACTTTTAAACAGCTAGCAGAACCCCTTAAACCGACACACCACTTCGATTTGT	4987
DB	637	TTCAAGAGACTTTTAAACAGCTAGCAGAACCCCTTAAACCGACACACCACTTCGATTTGT	578
QY	4988	TTAGCTATGATACAGGCTGAAATATAAACCCCGCACTAT 5025	
DB	577	TTAGCTATGATACAGGCTGAAATATAAGTCCGCTAT 540	

## RESULT 6

US-10-032-393-7/c  
 ; Sequence 7, Application US/10032393  
 ; Publication No. US20030027286A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Haselbeck, Robert  
 ; APPLICANT: Wall, Daniel  
 ; APPLICANT: Gross, Molly  
 ; TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE  
 ; FILE REFERENCE: ELITRA.010A  
 ; CURRENT APPLICATION NUMBER: US/10/032.393  
 ; CURRENT FILING DATE: 2001-12-21  
 ; PRIOR APPLICATION NUMBER: 60/259,434  
 ; PRIOR FILING DATE: 2000-12-27  
 ; PRIOR APPLICATION NUMBER: 09/948,993  
 ; PRIOR FILING DATE: 2001-09-06  
 ; PRIOR APPLICATION NUMBER: 60/230,335

```
; PRIOR FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 10929
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Vector pEP25
US-10-032-393-7

Query Match          20.2%; Score 1109.4; DB 15; Length 10929;
Best Local Similarity 99.9%; Pred. No. 2.8e-206;
Matches 1110; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      936 AAAAGAAAAACGAAATGATACACCATCAGTCGAAAATAAAGATATATATGGAGATTAAGAC 995
       |||
Db       5830 AAAAGAAAAACGAAATGATACACCAATCAGTCGAAAATAAAGATATATATGGAGATTAAGAC 5771
       |||
QY      996 GGTTCTGTGTTTCTGTGCTGACTTGCCACCATATCATATAAAATCGAAACAGCAAGAATGGCGG 1055
       |||
Db       5770 GGTTCTGTGTTTCTGTGCTGACTTGCCACCATATCATATAAAATCGAAACAGCAAGAATGGCGG 5711
       |||
QY      1056 AAACGTAATAAGAGAGTTATGGAATATAGACTTAGAACCAAACTTAAGAGTGTGTGATAGT 1115
       |||
Db       5710 AAACGTAATAAGAGAGTTATGGAATATAGACTTAGAACCAAACTTAAGAGTGTGTGATAGT 5651
       |||
QY      1116 GCAGTATCTTAAAAATTTTGTATATAGCAATTGAAGTTAAATTTAGATGCTTAAAAATTTGT 1175
       |||
Db       5650 GCAGTATCTTAAAAATTTTGTATATAGCAATTGAAGTTAAATTTAGATGCTTAAAAATTTGT 5591
       |||
QY      1176 AATTAGAAGGAGTGATTACATGAACAAAATATATAAATTTCTCAAACTTTTTTAACGA 1235
       |||
Db       5590 AATTAGAAGGAGTGATTACATGAACAAAATATATAAATTTCTCAAACTTTTTTAACGA 5531
       |||
QY      1236 GTCAAAAAGTACTCAACCAAATTAATAAACAAATTGAATTTTAAAGAAACCGATACCGTTT 1295
       |||
Db       5530 GTCAAAAAGTACTCAACCAAATTAATAAACAAATTGAATTTTAAAGAAACCGATACCGTTT 5471
       |||
QY      1296 ACGAATTTGAAACAGGTAAAGGGCATTTTACGACGAAACTGGCTAAATTAAGTAAACAGG 1355
       |||
Db       5470 ACGAATTTGAAACAGGTAAAGGGCATTTTACGACGAAACTGGCTAAATTAAGTAAACAGG 5411
       |||
QY      1356 TAACGTCATTATTGAATTAGACAGTCATCTATTCAACTTTATCGTCAGAAAAAATTTAAACTGA 1415
       |||
Db       5410 TAAAGTCATTATTGAATTAGACAGTCATCTATTCAACTTTATCGTCAGAAAAAATTTAAACTGA 5351
       |||
QY      1416 ATACTCGTGTCACTTTAATTCAACAAGATATCTACAGTTTCAATTCCTTAAACAAACAGA 1475
       |||
Db       5350 ATACTCGTGTCACTTTAATTCAACAAGATATCTACAGTTTCAATTCCTTAAACAAACAGA 5291
       |||
QY      1476 GGATATAAATTTGTTGGAGTATTCTTTACCATTTTAAAGCACACAAATTTATTAAAAAGTGG 1535
       |||
Db       5290 GGATATAAATTTGTTGGAGTATTCTTTACCATTTTAAAGCACACAAATTTATTAAAAAGTGG 5231
       |||
QY      1536 TTTTGAAGCCATCGGTCTGACATCTATCTGATTTGTGAAGAGGATTTCTCAAGCGGTA 1595
       |||
Db       5230 TTTTGAAGCCATCGGTCTGACATCTATCTGATTTGTGAAGAGGATTTCTCAAGCGGTA 5171
       |||
QY      1596 CTTTGGATATTCACCGAACACTAGGTTGCTCTTGCACACTCAAGTCTCGATTCAGCAAT 1655
       |||
Db       5170 CTTTGGATATTCACCGAACACTAGGTTGCTCTTGCACACTCAAGTCTCGATTCAGCAAT 5111
       |||
QY      1656 TGCTTAAGCTGCCAGGGAAATGCTTTTCATCCTTAAACCAAAGTAAACAGTGTCTTAATAA 1715
       |||
Db       5110 TGCTTAAGCTGCCAGGGAAATGCTTTTCATCCTTAAACCAAAGTAAACAGTGTCTTAATAA 5051
       |||
QY      1716 AACTTACCGCCATACACAGATGTTCCAGATAAATTTGGAGACTATATACACTCTTTC 1775
       |||
Db       5050 AACTTACCGCCATACACAGATGTTCCAGATAAATTTGGAGACTATATACACTCTTTC 4991
       |||
QY      1776 TTTTCAAAATGGSGTCAATTCGAAATATCGTCAACTGTTTACTTAAAAATTCAGTTTTCATCAAG 1835
```

Query Match	15.9%;	Score 876.6;	DB 15;	Length 9412;
Best Local Similarity	83.4%;	Pred. No. 9e-161;		
Matches 1102; Conservative	0;	Mismatches 9;	Indels 210;	Gaps 1;
Qy	936	AAAAGHAAAACGAAATGATACACCAATCAGTCGAAAAAAGAGATATAATGGAGATAAGAC	9395	
Db	6703	AAAAGHAAAACGAAATGATACACCAATCAGTCGAAAAAAGAGATATAATGGAGATAAGAC	6644	
Qy	996	GGTTCGTGTTCTGTCGTGACTTGCACCATATCATATAAAATCGAAAACAGCAAGAATCGCGG	1055	
Db	6643	GGTTCGTGTTCTGTCGTGACTTGCACCATATCATATAAAATCGAAAACAGCAAGAATCGCGG	6584	
Qy	1056	AAACGTAAAAGAAGTTATCGAAATTAAGACTTAAGACCAAACTTAAGAGTGTGTTGATAGT	1115	
Db	6583	AAACGTAAAAGAAGTTATCGAAATTAAGACTTAAGACCAAACTTAAGAGTGTGTTGATAGT	6524	
Qy	1116	GCAGTATCTTAAAAATTTTGATATAATAGGAATTTGAAGTTAAATTAGATGCTAAAAATTTGT	1175	
Db	6523	GCATTATCTTAAAAATTTTGATATAATAGGAATTTGAAGTTAAATTAGATGCTAAAAATTTGT	6464	
Qy	1176	AAATTAAGA-----	1183	
Db	6463	AAATTAAGAAGGGGNTTCGTGATTTGTTGATATCCAAATCGTAATGTAGATAAAACATC	6404	
Qy	1184	-----	1183	



```
QY 1386 TCAACTTATCGTCAGAAAATTAATACTGAATCTGCTGCTCACTTTAAATTCACCAAGATA 1445
Db 6043 TCAACTTATCGTCAGAAAATTAATACTGAATCTGCTGCTCACTTTAAATTCACCAAGATA 5984
QY 1446 TTCTACAGTTTCAATTCCTTAAACACAAACAGAGGTATATAATTTGTTGGAGTATTCCTTACC 1505
Db 5983 TTCTACAGTTTCAATTCCTTAAACACAAACAGAGGTATATAATTTGTTGGAGTATTCCTTACC 5924
QY 1506 ATTTAAGCACACAATTAATTAATAAGTGTGTTTGAAGCCATGCGTCTGACATCTATC 1565
Db 5923 ATTTAAGCACACAATTAATTAATAAGTGTGTTTGAAGCCATGCGTCTGACATCTATC 5864
QY 1566 TGATTGTTGAAGAGGATTCACAGGCTACCTTGGATATTCACCGAACACTAGGTTGC 1625
Db 5863 TGATTGTTGAAGAGGATTCACAGGCTACCTTGGATATTCACCGAACACTAGGTTGC 5804
QY 1626 TCTTGCACTCAAGTCTCGATTTCAGCAATTCCTTAAGCTGCGAGCGGAATGCTTTCATC 1685
Db 5803 TCTTGCACTCAAGTCTCGATTTCAGCAATTCCTTAAGCTGCGAGCGGAATGCTTTCATC 5744
QY 1686 CTAACCCAAAAGTAACAGTGTCTTAATAAATCTTACCCGCCATACCAAGATGTTCCAG 1745
Db 5743 CTAACCCAAAAGTAACAGTGTCTTAATAAATCTTACCCGCCATACCAAGATGTTCCAG 5684
QY 1746 ATAAATATTCGAAGCTATATACGTACTCTTTTCAAAATGGGTCAATTCGAGATATCGTC 1805
Db 5683 ATAAATATTCGAAGCTATATACGTACTCTTTTCAAAATGGGTCAATTCGAGATATCGTC 5624
QY 1806 AACTGTTTACTAAATCAGTTTCATCAGCAATGAACACGCCAAAGTAACAATTTAA 1865
Db 5623 AACTGTTTACTAAATCAGTTTCATCAGCAATGAACACGCCAAAGTAACAATTTAA 5564
QY 1866 GTACCGTTACTTATGAGCAAGTATGCTATATTTTAAATAGTATCTATTTAATTAAGCGGA 1925
Db 5563 GTACCGTTACTTATGAGCAAGTATGCTATATTTTAAATAGTATCTATTTAATTAAGCGGA 5504
QY 1926 GGAATAATTCATGAGTGCCTTTTGTAATTTGAAAGTTACAGTTTACTAAAGGGAAT 1985
Db 5503 GGAATAATTCATGAGTGCCTTTTGTAATTTGAAAGTTACAGTTTACTAAAGGGAAT 5444
QY 1986 GTAGATAAATTTAGTATCTACTGACAGCTTCCAAAGGAGTAAAGGTCCTTAGCG 2045
Db 5443 GGAGATAAATTTAGTATCTACTGACAGCTTCCAAAGGAGTAAAGGTCCTTAGCG 5384
QY 2046 C 2046
Db 5383 C 5383
```

```
RESULT 9
US-10-032-393-48/c
; Sequence 48, Application US/10032393
; Publication No. US20030027286A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Wall, Daniel
; APPLICANT: Gross, Molly
; TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE
; FILE REFERENCE: ELITRA.010A
; CURRENT APPLICATION NUMBER: US/10/032,393
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/259,434
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 9417
; TYPE: DNA
```

```
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Vector pPEPF14*
US-10-032-393-48
```

```
Query Match 15.9%; Score 876.6; DB 15; Length 9417;
Best Local Similarity 83.4%; Pred. No. 9e-161;
Matches 1102; Conservative 0; Mismatches 9; Indels 210; Gaps 1;

QY 936 AAAAGAAAACGAAATGATACACCAATCAGTGCAGAAAAAAGATATAATGGGAGATAAGAC 995
Db 6703 AAAAGAAAACGAAATGATACACCAATCAGTGCAGAAAAAAGATATAATGGGAGATAAGAC 6644
QY 996 GGTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1055
Db 6643 GGTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6584
QY 1056 AAACGTAAAGAAAGTTATGGAATTAAGACTTAGAAGCAAACTTAAGAGTGTGTTGATAGT 1115
Db 6583 AAACGTAAAGAAAGTTATGGAATTAAGACTTAGAAGCAAACTTAAGAGTGTGTTGATAGT 6524
QY 1116 GCAGTATCTTAAATTTTGTATATAAGGAATTAAGTTAAATTTAGATGCTTAAATTTGT 1175
Db 6523 GCATTTATCTTAAATTTTGTATATAAGGAATTAAGTTAAATTTAGATGCTTAAATTTGT 6464
QY 1176 AATTAAAGA----- 1183
Db 6463 AATTAAAGAAGGAGGATTGCTGTCATGTTGGTATTCCAAATGCGTAATGTAGATAAAACATC 6404
QY 1184----- 1183
Db 6403 TACTGTTTGAACACAGACTAAAAACAGTGAATAGCGAGATAATAATAATACCTTAGATTAA 6344
QY 1184----- 1183
Db 6343 TTCCTACAGTGACTAATCTTATGACTTTTAAACAGATAACTAAATTAACAACAAATC 6284
QY 1184----- 1205
Db 6283 GTTTAACTTCTGTTTGTGTTTATAGTGTATCACTTCAGGAGAGATTACATGAACAAA 6224
QY 1206 ATATAAATATTTCTCAAAACCTTTTAAAGAGTGAAGAGTCTCAACCAATATAATAAC 1265
Db 6223 ATATAAATATTTCTCAAAACCTTTTAAAGAGTGAAGAGTCTCAACCAATATAATAAC 6164
QY 1266 AATTGAATTTAAAGAAACCGATACCGTTTACGAAATTTGGAACAGGTAAAGGGCATTTAA 1325
Db 6163 AATTGAATTTAAAGAAACCGATACCGTTTACGAAATTTGGAACAGGTAAAGGGCATTTAA 6104
QY 1326 CGACAAACCTGGCTAAATAAGTAAACAGGTAAACAGTCTATTGAATTAAGACAGTCACTAT 1385
Db 6103 CGACAAACCTGGCTAAATAAGTAAACAGGTAAACAGTCTATTGAATTAAGACAGTCACTAT 6044
QY 1386 TCAACTTATCGTCAGAAAATTTAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1445
Db 6043 TCAACTTATCGTCAGAAAATTTAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5984
QY 1446 TTCTACAGTTTCAATTCCTTAAACACAAACAGAGGTATAATAATTTGTTGGAGTATTCCTTACC 1505
Db 5983 TTCTACAGTTTCAATTCCTTAAACACAAACAGAGGTATAATAATTTGTTGGAGTATTCCTTACC 5924
QY 1506 ATTTAAGCACACAATTAATTAATAAGTGTGTTTGAAGCCATGCGTCTGACATCTATC 1565
Db 5923 ATTTAAGCACACAATTAATTAATAAGTGTGTTTGAAGCCATGCGTCTGACATCTATC 5864
QY 1566 TGATTGTTGAAGAGGATTCACAGGCTACCTTGGATATTCACCGAACACTAGGTTGC 1625
Db 5863 TGATTGTTGAAGAGGATTCACAGGCTACCTTGGATATTCACCGAACACTAGGTTGC 5804
QY 1626 TCTTGCACTCAAGTCTCGATTTCAGCAATTCCTTAAGCTGCGAGCGGAATGCTTTCATC 1685
Db 5803 TCTTGCACTCAAGTCTCGATTTCAGCAATTCCTTAAGCTGCGAGCGGAATGCTTTCATC 5744
```









Db	6679	AAAAAGAAAAACGAATGATACACCAATCAGTCGCAAAAAAGATATATGCGAGATAAGAC	6620
Qy	996	GGTTTCGTGTTGCTGCTGACCTTGCACCATATCATATAAATCGAACAGCAAGATGCGG	1055
Db	6619	GGTTTCGTGTTGCTGCTGACCTTGCACCATATCATATAAATCGAACAGCAAGATGCGG	6560
Qy	1056	AAACGTAAGAAAGAGTTATGGAATAAGACTTAGAAGCAAACTTTAAGAGTGTGTGATAGT	1115
Db	6559	AAACGTAAGAAAGAGTTATGGAATAAGACTTAGAAGCAAACTTTAAGAGTGTGTGATAGT	6500
Qy	1116	GCAGTATCTTAAATTTTGTATATAGGAATGCAAGTTAAATTTAGATGCTAATAAATTTGT	1175
Db	6499	GCATTTATCTTAAATTTTGTATATAGGAATGCAAGTTAAATTTAGATGCTAATAAATTTGT	6440
Qy	1176	AATTAAGA-----	1183
Db	6439	AATTAAGAAGCGGATTGCTCATGTTGGTATTCCAAATGCGTAATGTAGATAAATCATC	6380
Qy	1184	-----	1183
Db	6379	TACTGTTTGAACAGACTAAAAACAGTGTANTACGCAGATAAATAAATACGTTAGATTAA	6320
Qy	1184	-----	1183
Db	6319	TTCTACCAGTGACTAATCTTATGACTTTTAAACAGATAAATAAATTACAAACAAATC	6260
Qy	1184	-----AGGAGTGATTACATGAACAAA	1205
Db	6259	GTTTAACTTCTGTATTTTATAGATGTAATCACTTCAGGAGAGATTACATGAACAAA	6200
Qy	1206	ATATAAAATTTCTCAAACTTTTAAACGAGTCAAAAGTACTCAACCAAAATAAATAAAC	1265
Db	6199	ATATAAAATTTCTCAAACTTTTAAACGAGTCAAAAGTACTCAACCAAAATAAATAAAC	6140
Qy	1266	AATTTGAATTTAAAGAAACCGATACCGTTTACGAAATTTGAAACAGGTAAGGCGCATTTAA	1325
Db	6139	AATTTGAATTTAAAGAAACCGATACCGTTTACGAAATTTGAAACAGGTAAGGCGCATTTAA	6080
Qy	1326	CGAGAAACTGGCTAAATAAGTAAACAGGTAAACGTTAATGAAATAGACAGTCACTAT	1385
Db	6079	CGAGAAACTGGCTAAATAAGTAAACAGGTAAACGTTAATGAAATAGACAGTCACTAT	6020
Qy	1386	TCAACTTATCGTCAGAAAAATTAATACTGCTGCTCACTTTAAATTCACCAAGATA	1445
Db	6019	TCACTTATCGTCAGAAAAATTAATACTGCTGCTCACTTTAAATTCACCAAGATA	5960
Qy	1446	TTCTACAGTTTCAATTCCTTAAACAAACAGAGGTATATAAATTTGTTGGAGTATTCCTTACC	1505
Db	5959	TTCTACAGTTTCAATTCCTTAAACAAACAGAGGTATATAAATTTGTTGGAGTATTCCTTACC	5900
Qy	1506	ATTAAAGCACACAAATTTATTAAGAAAGTGTGTTTGAAGCCCATGCTGTCATCATCTATC	1565
Db	5899	ATTAAAGCACACAAATTTATTAAGAAAGTGTGTTTGAAGCCCATGCTGTCATCATCTATC	5840
Qy	1566	TGATTTGTAAGAGGATTCTAAGCGTACCTTGGATATTCACGAAACACTAGGTTGTC	1625
Db	5839	TGATTTGTAAGAGGATTCTAAGCGTACCTTGGATATTCACGAAACACTAGGTTGTC	5780
Qy	1626	TCTTGCACTCAAGTCTCGATTCAGCAATTCGTTAAGCTGCCAGGGGATGCTTTCATC	1685
Db	5779	TCTTGCACTCAAGTCTCGATTCAGCAATTCGTTAAGCTGCCAGGGGATGCTTTCATC	5720
Qy	1686	CTAAACCAAAAGTAAACAGTGTCTTAAATAAATTTACCCGCCATACACAGATGTTCCAG	1745
Db	5719	CTAAACCAAAAGTAAACAGTGTCTTAAATAAATTTACCCGCCATACACAGATGTTCCAG	5660
Qy	1746	ATATAATTTGGAAGCTATATACGTACTTTGTTTCAAAATGGGTCAATTCAGAAATATCGTC	1805
Db	5659	ATATAATTTGGAAGCTATATACGTACTTTGTTTCAAAATGGGTCAATTCAGAAATATCGTC	5600
Qy	1806	AACGTGTTTACTAAAAATCAGTTTTCATCAAGCAATGAAACAGCGCAAGTAAACAAATTTAA	1865
Db	5599	AACGTGTTTACTAAAAATCAGTTTTCATCAAGCAATGAAACAGCGCAAGTAAACAAATTTAA	5540

RESULT 14

US-10-032-393-47/c  
; Sequence 47, Application US/10032393  
; Publication No. US20030027286A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Wall, Daniel  
; APPLICANT: Gross, Molly  
; TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE  
; FILE REFERENCE: ELITRA.010A  
; CURRENT APPLICATION NUMBER: US/10/032,393  
; CURRENT FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/259,434  
; PRIOR FILING DATE: 2000-12-27  
; PRIOR APPLICATION NUMBER: 09/948,993  
; PRIOR FILING DATE: 2001-09-06  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; NUMBER OF SEQ ID NOS: 68  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 47  
; LENGTH: 12733  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Vector pPEP14  
US-10-032-393-47

Query Match	15.9%;	Score 876.6;	DB 15;	Length 12733;
Best Local Similarity	83.4%;	Pred. No. 1.1e-160;		
Matches 1102;	Conservative 0;	Mismatches 9;	Indels 210;	Gaps 1;
Qy	936	AAAAGAAAAACGAATGATACACCAATCAGTGCAAAAAAGATATAATGGAGATAAGAC	995	
Db	12063	AAAAGAAAAACGAATGATACACCAATCAGTGCAAAAAAGATATAATGGAGATAAGAC	12004	
Qy	996	GGTTGCTGTTTCGTGCTGACTTGCACCATATCATATAAATCGAAACAGCAAGATGCGG	1055	
Db	12003	GGTTGCTGTTTCGTGCTGACTTGCACCATATCATATAAATCGAAACAGCAAGATGCGG	11944	
Qy	1056	AAACGTAAAGAAAGTATGGAATAAGACTTAGAAGCAAACTTAAGAGTGTGTTGATAGT	1115	
Db	11943	AAACGTAAAGAAAGTATGGAATAAGACTTAGAAGCAAACTTAAGAGTGTGTTGATAGT	11884	
Qy	1116	GCAGTATCTTAAAAATTTTGTATAATAGGAAATGAAGTTAAATTTAGATGCTAAAAATTTGT	1175	
Db	11883	GCATTTATCTTAAAAATTTTGTATAATAGGAAATGAAGTTAAATTTAGATGCTAAAAATTTGT	11824	
Qy	1176	AATTAAGA-----	1183	
Db	11823	AATTAAGAAGGAGGAGGATTCGTCATGTTGGTATTCCAAATCGTAATGTAGATAAATCATC	11764	
Qy	1184	-----	1183	
Db	11763	TACTGTTTGAACAGACTAAAAACAGTGATTACGCAGATAAATAAATACGTTAGATTAA	11704	



Qy	1386	TCAACTTATCGTCAGAAAAATTAAACTGAACTACGTCGTCACTTTAAATTCACCAAGATA	1445
Db	11409	TCAACTTATCGTCAGAAAAATTAAACTGAACTACGTCGTCACTTTAAATTCACCAAGATA	11350
Qy	1446	TTCTACAGTTTCAATTCCTCAACAAACAGAGGTATAAAAATTGTTGGAGTATTCCTTACC	1505
Db	11349	TTCTACAGTTTCAATTCCTCAACAAACAGAGGTATAAAAATTGTTGGAGTATTCCTTACC	11290
Qy	1506	ATTTAAGCACACAAATATTAATAAAGTGGTTTTGAAAGCCATGCGTCTGACATCTATC	1565
Db	11289	ATTTAAGCACACAAATATTAATAAAGTGGTTTTGAAAGCCATGCGTCTGACATCTATC	11230
Qy	1566	TGATTCGTTGAAGAGGATCTCAACGGTACCTTGGATATTCACCGAACACTAGGTTGC	1625
Db	11229	TGATTCGTTGAAGAGGATCTCAACGGTACCTTGGATATTCACCGAACACTAGGTTGC	11170
Qy	1626	TCCTGCACACTCAAGTCTCGAATTCAGCAATGCTTAAGCTGCCAGCGGAATGCTTTCATC	1685
Db	11169	TCCTGCACACTCAAGTCTCGAATTCAGCAATGCTTAAGCTGCCAGCGGAATGCTTTCATC	11110
Qy	1686	CTAAACCAAAGTAAACAGTGTCTTAATAAAGTACCCGCCATACACAGATGTTCCAG	1745
Db	11109	CTAAACCAAAGTAAACAGTGTCTTAATAAAGTACCCGCCATACACAGATGTTCCAG	11050
Qy	1746	ATAAATATTCGAAAGCTATATACGTAATGTTGTTTCAAAATGGGTCAATCGAGAATATCGTC	1805
Db	11049	ATAAATATTCGAAAGCTATATACGTAATGTTGTTTCAAAATGGGTCAATCGAGAATATCGTC	10990
Qy	1806	AACTGTTTACTAAAAATCAGTTTTCATCAAGCAATGAAACACGCCAAAGTAAACAATTTAA	1865
Db	10989	AACTGTTTACTAAAAATCAGTTTTCATCAAGCAATGAAACACGCCAAAGTAAACAATTTAA	10930
Qy	1866	GTACCGTTACTTATGAGCAAGTATGTTCTATTTTAAATAGTTATCTATTATTAAACGGGA	1925
Db	10929	GTACCAATTAATGAGCAAGTATGTTCTATTTTAAATAGTTATCTATTATTAAACGGGA	10870
Qy	1926	GGAAATAAATCTATGAGTCGTTTTGTTAAATTTGAAAGTTACACGTTACTAAAGGGAAT	1985
Db	10869	GGAAATAAATCTATGAGTCGTTTTGTTAAATTTGAAAGTTACACGTTACTAAAGGGAAT	10810
Qy	1986	GTAGATAAATTAATGATATCTACTGACAGCTTCCAGAGCTAAAGAGTCCCTAGCG	2045
Db	10809	GGAGATAAATTAATGATATCTACTGACAGCTTCCAGAGCTAAAGAGTCCCTAGCG	10750
Qy	2046	C 2046	
Db	10749	C 10749	

Search completed: September 2, 2004, 06:05:50  
Job time : 2788 secs

This Page Blank (uspro)



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: September 1, 2004, 14:20:49 : Search time 12630 Seconds  
(without alignments)  
12997.023 Million cell updates/sec

Title: US-10-030-390-2  
Perfect score: 5497  
Sequence: 1 gaattcgattagtcattctt.....ctagtcgaagtatttaata 5497

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues  
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*\*

- 1: em\_estba:\*\*
- 2: em\_estbun:\*\*
- 3: em\_estin:\*\*
- 4: em\_estmu:\*\*
- 5: em\_estov:\*\*
- 6: em\_estpl:\*\*
- 7: em\_estro:\*\*
- 8: em\_htc:\*\*
- 9: gb\_esti:\*\*
- 10: gb\_est2:\*\*
- 11: gb\_htc:\*\*
- 12: gb\_est3:\*\*
- 13: gb\_est4:\*\*
- 14: gb\_est5:\*\*
- 15: em\_estfun:\*\*
- 16: em\_estom:\*\*
- 17: em\_gss\_hum:\*\*
- 18: em\_gss\_inv:\*\*
- 19: em\_gss\_pln:\*\*
- 20: em\_gss\_vrt:\*\*
- 21: em\_gss\_fun:\*\*
- 22: em\_gss\_man:\*\*
- 23: em\_gss\_mus:\*\*
- 24: em\_gss\_pro:\*\*
- 25: em\_gss\_rod:\*\*
- 26: em\_gss\_pbg:\*\*
- 27: em\_gss\_vri:\*\*
- 28: gb\_gss1:\*\*
- 29: gb\_gss2:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	204	3.7	354	13	BY102466
2	204	3.7	390	9	AI323443
3	204	3.7	451	9	AI893383
4	204	3.7	455	9	AI323089

5	204	3.7	456	9	AA060105
6	204	3.7	456	14	W83072
7	204	3.7	473	13	BY708773
8	204	3.7	474	11	AK008625
9	204	3.7	474	11	AK010179
10	204	3.7	475	11	AK008809
11	204	3.7	477	11	AK008873
12	204	3.7	486	13	BY708705
13	204	3.7	505	13	BY708809
14	204	3.7	505	9	AV072530
15	200.8	3.7	384	13	BY076652
16	200.4	3.6	354	13	BY101861
17	191.2	3.5	311	10	BB564851
18	188	3.4	435	9	AA097254
19	186.4	3.4	311	10	BB564891
20	181.8	3.3	443	14	W12684
21	168.4	3.1	307	10	BB565090
22	166.6	3.0	306	10	BB565053
23	165	3.0	238	29	AY419513
24	160.2	2.9	263	10	BB564847
25	156.4	2.8	332	9	AV076259
26	155	2.8	705	10	BB667664
27	155	2.8	1515	11	AK050258
28	153.8	2.8	300	10	BB564905
29	141	2.6	421	10	BB738530
30	139.4	2.5	1658	28	BH770822
31	138	2.5	354	10	BF707544
32	136.2	2.5	298	10	BB565023
33	135.6	2.5	264	10	BB564882
34	132.8	2.4	242	10	BB565149
35	130.8	2.4	812	12	BM536445
36	130.6	2.4	255	10	BB564787
37	129.8	2.4	602	28	B06973
38	125.8	2.3	254	10	BB565113
39	124.8	2.3	317	9	AV075620
40	122.6	2.2	320	9	AV081837
41	119.8	2.2	304	9	AV075572
42	116.6	2.1	301	9	AV077679
43	116.2	2.1	208	10	BB564887
44	116.2	2.1	308	9	AV080925
45	115.8	2.1	300	9	AV074919

ALIGNMENTS

RESULT 1

BY102466 354 bp mRNA linear EST 07-DEC-2002  
LOCUS BY102466 RIKEN full-length enriched, pooled tissues, adult spleen,  
DEFINITION etc. Mus musculus cDNA clone K630145D13 5', mRNA sequence.  
ACCESSION BY102466  
VERSION BY102466.1 GI:26213083  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,  
Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamakawa, T.,  
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,  
Schonbach, C., Gofobori, T., Baldarelli, R., Hill, D.P., Bult, C.,  
Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,  
Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V.,  
Chothia, C., Corbani, L.E., Cousins, S., Balla, E., Dragani, T.A.,  
Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,  
Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,  
Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, B.D., Kanai, A.,  
Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Kongsawa, A.,  
Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,  
Maitais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,  
Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G.,





FEATURES	Location/Qualifiers	
	source	1..455
FEATURES	Location/Qualifiers	
	source	1..455
MGI:292325		
Putative full length read		
vector to vector length is 457		
Seq primer: -28M13 rev2 from Amersham.		
Location/Qualifiers		
1..455		
/organism="Mus musculus"		
/mol_type="mRNA"		
/db_xref="taxon:10090"		
/clones="IMAGE:481581"		
/dev_stage="19.5 dpc total fetus"		
/lab_host="DH10B (ampicillin resistant)"		
/clone_lib="Soares mouse p3NMF19.5"		
/notes="vector: p7713D (Pharmacia) with a modified		
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA		
was primed with a Not I - oligo(dT) primer [5,		
TGTACCAATCTGAAGTGGAGCGCGCATTTTTTTTTTTT 3'],		
double-stranded cDNA was size selected, ligated to Eco RI		
adapters (Pharmacia), digested with Not I and cloned into		
the Not I and Eco RI sites of a modified p7713 vector		
(Pharmacia). Library went through one round of		
normalization to a Cot = 5. Library constructed by Bento		
Soares and M.Patima Bonaldo. RNA was kindly provided by		
Dr. Minoru Ko (Wayne State University)."		
ORIGIN		
Query Match 3.7%; Score 204; DB 9; Length 455;		
Best Local Similarity 100.0%; Pred. No. 5.6e-26;		
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	296	GCCCAGGCCAGGCCAGGCCAGGCCAGGAGAAACATGTATCATGCGCCCGGGAG 355
DB	73	GCCCAGGCCAGGCCAGGCCAGGCCAGGAGAAACATGTATCATGCGCCCGGGAG 132
QY	356	AGGATAATTGCTTCCCGGTGTCACCGCCAGCAGTGCACGGAGAGAGTGTCTGT 415
DB	133	AGGATAATTGCTTCCCGGTGTCACCGCCAGCAGTGCACGGAGAGAGTGTCTGT 422
QY	416	TTTGATGACAGTGTCCCGGGATTCCCGTGTGCTTCCACCCCATGGCCATCGAGAACT 475
DB	193	TTTGATGACAGTGTCCCGGGATTCCCGTGTGCTTCCACCCCATGGCCATCGAGAACT 252
QY	476	CAAGAAGAAGATGTCCTTCTAA 499
DB	253	CAAGAAGAAGATGTCCTTCTAA 276
RESULT 5		
AA060105		
LOCUS		
DEFINITION		
m77111.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone		
IMAGE:481581 5' similar to gb:221858 M.Musculus mRNA for P domain		
protein (MOUSE); mRNA sequence.		
ACCESSION		
AA060105		
VERSION		
AA060105.1 GI:1553793		
KEYWORDS		
EST.		
SOURCE		
Mus musculus (house mouse)		
ORGANISM		
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
1 (bases 1 to 456)		
REFERENCE		
AUTHORS		
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,		
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,		
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,		
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and		
Waterston, R.		
The WashU-HMI Mouse EST Project		
Unpublished (1996)		
Contact: Marra M/Mouse EST Project		
WashU-HMI Mouse EST Project		
Washington University School of MedicineP		
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108		
Tel: 314 286 1800		
Fax: 314 286 1810		
Email: mouseest@watson.wustl.edu		
This clone is available royalty-free through LNL; contact the		
IMAGE Consortium (info@image.lnl.gov) for further information.		
TITLE		
JOURNAL		
COMMENT		
Unpublished (1996)		
Contact: Marra M/Mouse EST Project		
WashU-HMI Mouse EST Project		
Washington University School of MedicineP		
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108		
Tel: 314 286 1800		
Fax: 314 286 1810		
Email: mouseest@watson.wustl.edu		
This clone is available royalty-free through LNL; contact the		
IMAGE Consortium (info@image.lnl.gov) for further information.		

**Contact:** Yoshihide Hayashizaki  
Laboratory for genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-3222  
Fax: 81-45-503-3216

Fax: 81-45-503-9216  
 E-mail: genome-res@gsc.riken.go.jp,  
 URL: <http://genome.gsc.riken.go.jp/>  
 Aizawa, K., Akimura, T., Arakawa, T., Carninci, P.,  
 Adachi, J., Aizawa, K., Hirozane, T., Hori, F.,  
 Fukuda, S., Hashizume, W., Hayashida, K.,  
 Imetani, K., Ishii, Y., Itoh, M., Kagawa, I.,  
 Kawai, J., Kojima, Y., Kondo, S.,  
 Konno, H., Koya, S., Miyazaki, A., Murata, M.,  
 Nakamura, M., Nomura, K., Numazaki, R., Ono, M.,  
 Obsato, N., Saito, R., Sakazume, N.,  
 Sano, H., Sasaki, D., Sato, K., Shibata, K.,  
 Shiraki, T., Tgami, M., Tanaka, Y.,  
 Waki, K., Watanishi, A., Muramatsu, M.,  
 and Hayashizaki, Y.  
 Direct Submission  
 Computational Analysis of Full-length Mouse cDNAs Compared with  
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 Genes. Genome Res. 10 (10), 1617-1630 (2000)  
 RIKEN integrated sequence analysis (RISA) system-384-format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
 10 (11), 1757-1771 (2000)

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

Location/Qualifiers  
1. .473

```

/organism="mus_musculus"
/mol_type="mrna"
/db_xref="taxon:10090"
/clone="2210403105"
/sex="male"
/tissue_type="stomach"
/dev_stage="adult"
/lab_host="SOLZ"
/clone_lib="RIKEN full-length enriched, adult male
stomach"
/notes="Site 1: XhoI; Site 2: SmaI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Institute, Wako, Japan."

```





Hiraka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saico, R., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, I., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muranatsu, M. and Hayaeshizaki, Y.

Direct Submission  
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)  
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5', GAGAGAGAGAGATCCAGAGCTCTTTTTTTTTTNN 3'], cDNA was prepared by using thermostable thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adaptor of sequence [5', GAGAGAGATCTCGAGTATTAATTAATCACTCCCTCCCTCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.

Host: XhoI.

Location/Qualifiers  
1. 474  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="CS7BL/6J"  
/db\_xref="FANTOM DB:2310075107"  
/db\_xref="MG1:1898578"  
/db\_xref="taxon:10090"  
/clone="2310075107"  
/sex="male"  
/tissue\_type="tongue"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stages="adult"  
31..294  
/note="unnamed protein product; putative  
trefol factor 1 (MGD|MG1:89135)"  
/codon\_start=1  
/protein\_id="BAB26751.1"  
/db\_xref="GI:12845438"  
/translation="MEHKVICVLAIVLAFGLSLAQAOAQAQAEETCIMAPEPRINC  
GFGPGVTAQCCTERGCFDSVRGFPWCFPHPAIENTQEBCPF"

ORIGIN

Query Match	3.7%	Score 204;	DB 11;	Length 474;
Best Local Similarity	100.0%	Pred. No. 5.5e-26;		
Matches 204;	Conservative	0;	Mismatches 0;	Indels 0; Gaps 0;

296	GCCAGGCCAGGCCAGGCCAGGCCAGGAAACATGTATCATGGCCCCCGGGAG	355
91	GCCAGGCCAGGCCAGGCCAGGCCAGGAAACATGTATCATGGCCCCCGGGAG	150
356	AGGATAATTGTGGCTTCCCGGTGTACCGCCAGCAGTGCACGGAGAGAGTTGCTGT	415
151	AGGATAATTGTGGCTTCCCGGTGTACCGCCAGCAGTGCACGGAGAGAGTTGCTGT	210
416	TTTGATGACAGTGTCCGGGGATTCCGPGGTGCTTCCACCCCATGGCCATCGAGAACACT	475
211	TTTGATGACAGTGTCCGGGGATTCCGPGGTGCTTCCACCCCATGGCCATCGAGAACACT	270
476	CAAGAAGAAGATGTCCCTTTCTAA	499
271	CAAGAAGAAGATGTCCCTTTCTAA	294

RESULT 10	AK008809	LOCUS	AK008809	475 bp	mRNA	linear	HTC 20-SEP-2003
DEFINITION	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
ACCESSION	AK008809						
VERSION	AK008809.1						
KEYWORDS	HTC; CAP trapper.						
SOURCE	Mus musculus (house mouse)						
ORGANISM	Mus musculus						
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.						
1	Carninci, P. and Hayashizaki, Y.						
AUTHORS	High-efficiency full-length cDNA cloning						
TITLE	Mech. Enzymol. 303, 19-44 (1999)						
JOURNAL	9279253						
MEDLINE	10349636						
PUBMED							
REFERENCE	2						
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.						
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes						
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)						
MEDLINE	20499374						
PUBMED	11042159						
REFERENCE	3						
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsu, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Kashiwagi, K., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuo, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.						
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer						
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)						
MEDLINE	20530913						
PUBMED	11076861						
REFERENCE	4						
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.						
TITLE	Functional annotation of a full-length mouse cDNA collection						
JOURNAL	Nature 409, 685-690 (2001)						
REFERENCE	5						
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.						
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs						
JOURNAL	Nature 420, 563-573 (2002)						
REFERENCE	6 (bases 1 to 475)						
AUTHORS	Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, P., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toyota, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.						
TITLE	Direct Submission						
JOURNAL	Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@genomics.go.jp). URL: http://genome.gsc.riken.go.jp/. Tel: 81-45-503-9222.						
REFERENCE	7						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:221040321 product:trefoil factor 1, full insert sequence.						
JOURNAL	AK008809						
REFERENCE	AK008809						
AUTHORS	AK008809						
TITLE</							



Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawai, H., Kawasawa, Y., Kiedziarski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wyshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Nakazawa, M., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
 Nature 420, 563-573 (2002)  
 22354683  
 12466851  
 Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@gsr.riken.go.jp/  
 URL: http://genome.gsc.riken.go.jp/  
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y., Komuro, S., Konno, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission  
 Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

Location/Qualifiers  
 1. 486  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="221040321"  
 /sex="male"  
 /tissue type="stomach"  
 /dev stage="adult"  
 /lab\_host="SOLR"  
 /clone\_lib="RIKEN full-length enriched, adult male stomach"  
 /note="Site 1: XhoI; Site 2: SstI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'-GAGAGAGAGCGCCGCAACTGAGTTTCTTTTCTTTTNN 3']. cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGAGAGATCTTCGAGTAAATTAATCCCTCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI."

ORIGIN	Query Match	3.7%	Score 204	DB 13	Length 486
	Best Local Similarity	100.0%	Pred. No. 5.5e-26		
	Matches 204	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	296	CCCCAGGCCCGCCAGGCCCGCCAGGAGGAAACATGATCATGCGCCCGGGAG	355		
DB	91	CCCCAGGCCCGCCAGGCCCGCCAGGAGGAAACATGATCATGCGCCCGGGAG	150		
QY	356	AGGTAATATTCGCTTCCCGGTGTCACCCCGAGGAGTGCACGAGAGAGGTTGCTGT	415		
DB	151	AGGTAATATTCGCTTCCCGGTGTCACCCCGAGGAGTGCACGAGAGAGGTTGCTGT	210		
QY	416	TTTGATGACAGTGTCCGGGGATTCGGTGGTTCACCCCGATGGCATCGAGAACT	475		
DB	211	TTTGATGACAGTGTCCGGGGATTCGGTGGTTCACCCCGATGGCATCGAGAACT	270		
QY	476	CAAGAGAGAGATGTCCTTCTAA	499		
DB	271	CAAGAGAGAGATGTCCTTCTAA	294		

RESULT 13  
 BY708809  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 REFERENCE  
 AUTHORS

BY708809 505 bp mRNA linear EST 16-DEC-2002  
 musculus cDNA clone 2210414A03 5', mRNA sequence.  
 BY708809  
 BY708809.1 GI:27120003  
 EST  
 Mus musculus (house mouse)  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 505)  
 Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaide, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schrim, L.M., Kanapin, A., Matsuda, H., Batalov, S., Baesel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gusninch, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawajiri, H., Kawasawa, Y., Kiedziarski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wyshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Nakazawa, M., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,

Rogers,J., Birney,E. and Hayashizaki,Y.  
 Analysis of the mouse transcriptome based on functional annotation  
 of 60,770 full-length cDNAs  
 Nature 420, 563-573 (2002)  
 22354683  
 MEDLINE  
 PUBMED  
 COMMENT

Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center(GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@gsc.riken.go.jp,  
 URL:<http://genome.gsc.riken.go.jp/>  
 Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carminci,P.,  
 Fukuda,S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F.,  
 Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y.,  
 Kondo,S., Konno,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M.,  
 Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Saito,R., Sakazume,N.,  
 Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M.,  
 Takeda,Y., Waki,K., Wataniki,A., Muramatsu,M. and Hayashizaki,Y.  
 Direct Submission  
 Computational Analysis of Full-Length Mouse cDNAs Compared with  
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. Genome Res. 10 (10), 1617-1630 (2000)  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
 10 (11), 1757-1771 (2000)  
 Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 cDNA library was prepared and sequenced in Mouse Genome  
 Encyclopedia Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
 Division of Experimental Animal Research in Riken contributed to  
 prepare mouse tissues.  
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
 further details.

FEATURES  
 source

Location/Qualifiers  
 1..505  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="221041403"  
 /sex="male"  
 /tissue\_type="stomach"  
 /dev\_stage="adult"  
 /lab\_host="SOLR"  
 /clone\_lib="RIKEN full-length enriched, adult male  
 stomach"  
 /note="Site 1: XhoI, Site 2: SstI; cDNA library was  
 prepared and sequenced in Mouse Genome Encyclopedia  
 Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in  
 RIKEN. Division of Experimental Animal Research in Riken  
 contributed to prepare mouse tissues. 1st strand cDNA was  
 primed with a primer [5'  
 GAGAGAGAGCGCGCAACTCGAGTTTCTTTTCTTTT 3'], cDNA was  
 prepared by using reverse thermo-activated reverse  
 transcriptase and subsequently enriched for full-length by  
 cap-trapper. Second strand cDNA was prepared with the  
 primer adapter of sequence [5'  
 GAGAGAGAGTTCTCGAGTAATTAATTAATTC  
 was cleaved with XhoI and SstI."

## ORIGIN

```

Query Match          3.7%; Score 204; DB 13; Length 505;
Best Local Similarity 100.0%; Pred. No. 5.4e-26;
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

[illegible]

RESULT 14	TITLE
AV072530	JOURNAL
LOCUS	COMMENT
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	

AV072530 994 bp mRNA linear EST 24-OCT-2001  
AV072530 Mus musculus stomach C57BL/6J adult Mus musculus cDNA  
clone 2200004B10, mRNA sequence.

AV072530  
AV072530.2 GI:16380964

EST.  
Mus musculus (house mouse)  
Mus musculus

REFERENCE  
AUTHORS

Hata, A., Hiramoto, K., Horii, F., Ishii, Y., Ito, M., Kawai, J., Kikuchi, T., Kikunaga, T., Fukuda, S., Furuno, W., Hanganaki, I., Kato, A., Kato, K., Koyu, S., Matsuyama, T., Miyazaki, A., Nomura, Kono, H., Kouda, M., Okada, T., Saito, R., Sakai, C., Sakai, K., Onno, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toyota, T., Muramatsu, M. and Hayashizaki, Y. RIKEN Mouse ESTs (Arakawa, T., et al. 2001).

**JOURNAL  
COMMENT**

On Jun 24, 1999 this sequence version replaced qi:5192358.

Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suchoiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@gsc.riken.go.jp,  
 URL: <http://genome.gsc.riken.go.jp/>

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-tripped-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. *Genome Res.* 10 (10), 1617-1630 (2000).

wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, K., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsushima, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci,

Sugahara, Y. and Hayashizaki, Y.  
Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)

Kondo, S., Shingawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Alzawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and Hayashizaki, Y.

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

# FEATURES

source  
1. .994  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="P2200004810"  
/sex="male"  
/tissue\_type="stomach"  
/dev\_stage="adult"  
/clone\_lib="Mus musculus stomach C57BL/6J adult"

## ORIGIN

Query Match 3.7%; Score 204; DB 9; Length 994;  
Best Local Similarity 100.0%; Pred. No. 4.4e-26;  
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 296 GCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGAAACATGTATCATGGCCCCCGGAG 355  
Db 610 GCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGAAACATGTATCATGGCCCCCGGAG 669  
Qy 356 AGGATAAATTGGCTTCCCGGTGTCACCGCCAGCAGTCACGGAGAGAGGTTGCTGT 415  
Db 670 AGGATAAATTGGCTTCCCGGTGTCACCGCCAGCAGTCACGGAGAGAGGTTGCTGT 729  
Qy 416 TTGTATGACAGTGTCCGGGATTCCTCGTGGTCTTCCACCCATGGCCATCAGAACT 475  
Db 730 TTGTATGACAGTGTCCGGGATTCCTCGTGGTCTTCCACCCATGGCCATCAGAACT 789  
Qy 476 CAAGAGAGAGATGCCCTTCTAA 499  
Db 790 CAAGAGAGAGATGCCCTTCTAA 813

## RESULT 15

BY076652 RIKEN full-length enriched, pooled tissues, adult spleen,  
LOCUS EST. 384 bp mRNA linear EST 06-DEC-2002  
DEFINITION etc. Mus musculus cDNA clone K630005115 5', mRNA sequence.

## ACCESSION

VERSION BY076652  
KEYWORDS Mus musculus (house mouse)  
SOURCE EST.  
ORGANISM Mus musculus

## REFERENCE

AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 384)  
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaide, I., Oshio, N., Saito, R., Suzuki, H., Yananaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C., Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K. W., Blake, J. A., Bradt, D., Bruscia, V., Chothia, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grillo, M., S., Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Leshard, B., Lyons, P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S., Ravasi, T., Read, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C. A., Setou, M., Shimada, M., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavalan, M., Zhu, Y., Zimmer, A., Carninci, P.,

Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shitaki, T., Waki, K., Kawai, J., Aizawa, K., Aizawa, T., Fukuda, S., Hata, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S., Rogers, J., Birney, E. and Hayashizaki, Y.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)  
22354683  
12466851

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## COMMENT

Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gsc.riken.go.jp,  
URL: http://genome.sgc.riken.go.jp/

Aizawa, K., Akimura, T., Arakawa, I., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission  
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

## FEATURES

### source

Location/Qualifiers  
1. .384  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="K630005115"  
/clone\_lib="RIKEN full-length enriched, pooled tissues,  
adult spleen, etc."

/note="(dev\_stage=adult,tissue\_type=spleen,sex=male),  
(dev\_stage=adult,tissue\_type=kidney,sex=male),  
(dev\_stage=adult,tissue\_type=testis,sex=male),  
(dev\_stage=adult,tissue\_type=thymus,sex=male),  
(dev\_stage=adult,tissue\_type=heart,sex=male),  
(dev\_stage=adult,tissue\_type=colon,sex=male),  
(dev\_stage=adult,tissue\_type=stomach,sex=male),  
(dev\_stage=adult,tissue\_type=liver,sex=male),  
(dev\_stage=13 days embryo,tissue\_type=whole  
body,sex=mix), (dev\_stage=14 days embryo,tissue\_type=whole  
body,sex=mix), (dev\_stage=16 days embryo,tissue\_type=whole  
body,sex=mix), (dev\_stage=17 days embryo,tissue\_type=whole  
body,sex=mix), (dev\_stage=15 days pregnant,  
adult,tissue\_type=amion,sex=female), (dev\_stage=10 days  
neonate,tissue\_type=brain,sex=mix), (dev\_stage=10 days  
neonate,tissue\_type=thymus,sex=mix), (dev\_stage=10 days  
neonate,tissue\_type=heart,sex=mix)"

## ORIGIN

Query Match 3.7%; Score 200.8; DB 13; Length 384;  
Best Local Similarity 99.0%; Pred. No. 2.2e-25;



Matches	202;	Conservative	0;	Mismatches	2;	Indels	0;	Gaps	0;
Qy	296	GCC	CAG	CCC	CAG	CCC	CAG	CCC	CAG
Db	87	GCC	CAG	CCC	CAG	CCC	CAG	CCC	CAG
Qy	356	AGG	ATA	AAT	TGT	GCT	TCC	CGG	GAT
Db	147	AGG	ATA	AAT	TGT	GCT	TCC	CGG	GAT
Qy	416	TTT	GAT	GAC	AGT	TCC	CGG	GAT	TCC
Db	207	TTT	GAT	GAC	AGT	TCC	CGG	GAT	TCC
Qy	476	CAA	GAA	GAA	GAT	GTC	CTT	CTT	TAA
Db	267	CAA	GAA	GAA	GAT	GTC	CTT	CTT	TAA

Search completed: September 2, 2004, 00:11:16  
 Job time : 12633 secs

This Page Blank (usp'o)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 1, 2004, 13:38:44 ; Search time 20435 Seconds  
(without alignments)  
11659.244 Million cell updates/sec

Title: US-10-030-390-2  
Perfect score: 5497  
Sequence: 1 gaatcgattagtcatttt.....ctagtcgaagtgattaaata 5497

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues  
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl :  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.roi.\*  
11: gb.sts.\*  
12: gb.sts.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*  
29: em.vi.\*  
30: em.htg.hum.\*  
31: em.htg.inv.\*  
32: em.htg.other.\*  
33: em.htg.mus.\*  
34: em.htg.pln.\*  
35: em.htg.rod.\*  
36: em.htg.mam.\*  
37: em.htg.vrt.\*  
38: em.sy.\*  
39: em.hgo.hum.\*  
40: em.hgo.mus.\*  
41: em.hgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	DB ID	Description
1	5479	99.7	5505	6	AX069290 Sequence
2	4994.4	90.9	5217	6	AR381898 Sequence
3	4994.4	90.9	5217	6	AX001271 Sequence
4	4993.4	90.8	5230	6	AR381899 Sequence
5	4993.4	90.8	5230	6	AX001272 Sequence
6	4944.6	90.0	5231	6	AR381900 Sequence
7	4944.6	90.0	5231	6	AX001273 Sequence
8	4025.2	73.2	8072	12	AF038666 Cloning v
9	3816.6	69.4	8353	12	AY303239 Cloning v
10	2908.4	52.9	8199	1	AF007787 Enterococ
11	2816	51.2	4963	12	AF041239 Cloning v
12	2814.4	51.2	4698	12	AF039139 Cloning v
13	2807.6	51.1	11925	12	AY303237 Shuttle v
14	2807.6	51.1	13515	12	AY303236 Shuttle v
15	1735.4	31.6	9084	1	SPPLASS X64695 S.pyogenes
16	1735.4	31.6	14928	12	AY230218 Expressio
17	1735.4	31.6	19203	1	SPPLASMD X66468 S.pyogenes
18	1735.4	31.6	28975	1	AY357120 Streptoco
19	1735.4	31.6	28975	1	AY357120 Streptoco
20	1721	31.3	50237	1	EFCA501 X92945 Enterococcu
21	1720	31.3	6437	1	SAU83488 U83488 Streptococc
22	1678	30.5	11827	12	SVE509853 Cloning v
23	1645.2	29.9	19203	1	SPPLASMD X66468 S.pyogenes
24	1579.8	28.7	2270	1	SAPREPR X17655 Streptococc
25	1579.8	28.7	5842	1	AACOP X72021 S.agalactia
26	1461.4	26.6	66320	1	AE016833 Enterococ
27	1434.8	26.1	3765	6	BD193693 Enterococ
28	1417.4	25.8	1491	1	SPREPS X16803 Streptococc
29	1117.2	20.3	5019	12	CVE243541 Cloning v
30	1116	20.3	4523	12	EVE243540 Expressio
31	1115	20.3	8774	12	IVU69267 Integratio
32	1115	20.3	9065	12	AF405698 Reporter
33	1115	20.3	12168	12	AF405696 Reporter
34	1115	20.3	12248	12	AF405697 Reporter
35	1113.4	20.3	8611	12	AF072806 Integrati
36	1112	20.2	1126	1	SFAMB1 Y00116 Streptococc
37	1111	20.2	2833	1	CDI294529 Clostridi
38	1111	20.2	3410	12	AF435436 Cloning v
39	1111	20.2	10323	6	BD244511 Modificat
40	1111	20.2	10323	6	AR267509 Sequence
41	1111	20.2	17111	6	BD251969 Artificial
42	1111	20.2	17116	6	BD251970 Artificial
43	1110	20.2	6550	12	SYNPH2515 D90381 Synthetic c
44	1109.4	20.2	4129	1	CPUI8931 ui8931 clostridium
45	1109.4	20.2	9071	12	AB057644 Shuttle v

ALIGNMENTS

RESULT 1  
AX069290  
LOCUS AX069290 Sequence 2 from Patent WO0102570.  
DEFINITION AX069290  
ACCESSION AX069290  
VERSION AX069290.1 GI:12579163  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM artificial sequences.  
REFERENCE  
1  
AUTHORS Hans,W.C., Steidler,L. and Remaut,E.R.  
TITLE Delivery of trefol peptides  
JOURNAL Patent: WO 0102570-A 2 11-JAN-2001;  
Vlaams Interuniversitair Instituut voor Biotechnologie (BE)

linear PAT 25-JAN-2001

FEATURES		Location/Qualifiers	
source		1..5505	
		/organism="synthetic construct"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:32630"	
		/notes="Synthetic construct : plasmid"	
ORIGIN			
	Query Match	99.7%; Score 5479; DB 6; Length 5505;	
	Best Local Similarity	99.9%; Pred. No. 0;	
	Matches 5497; Conservative	0; Mismatches 0; Indels 8; Gaps 1;	
QY	1	GAATTCGATTAGTCACTCTCTTTTATTAGTTTCTTATATCTTAATGATAACA	60
DB	1	GAATTCGATTAGTCACTCTCTTTTATTAGTTTCTTATATCTTAATGATAACA	60
QY	61	TTTTTATAATTAATCTATAAACCATATCCCTCTTTTGAATCAAAATTTATTTACTCTCC	120
DB	61	TTTTTATAATTAATCTATAAACCATATCCCTCTTTTGAATCAAAATTTATTTACTCTCC	120
QY	121	TTTGTAGATATGTTATAATACAGTATCAGATCTGGGAGACCAACACGGTTTCCCACTAG	180
DB	121	TTTGTAGATATGTTATAATACAGTATCAGATCTGGGAGACCAACACGGTTTCCCACTAG	180
QY	181	AAATAATTTTGTTPAACCTTTAGAAAGGAGATATAGCATGAAAAAAGATTATCTCAGC	240
DB	181	AAATAATTTTGTTPAACCTTTAGAAAGGAGATATAGCATGAAAAAAGATTATCTCAGC	240
QY	241	TATTTTAATGCTACAGTCACTTTCTGCTCAGCCCGCTGTCAGGTGTTTACGCCCA	300
DB	241	TATTTTAATGCTACAGTCACTTTCTGCTCAGCCCGCTGTCAGGTGTTTACGCCCA	300
QY	301	GGCCAGGCCCGCAGGCCAGGAGGAAACATGATCATGGCCCGCCCGGAGAGGAT	360
DB	301	GGCCAGGCCCGCAGGCCAGGAGGAAACATGATCATGGCCCGCCCGGAGAGGAT	360
QY	361	AAATTGTGGCTTCCCGGTTGTCACCGCCAGCAGTGCACGGAGAGAGTTGCTGTTTGA	420
DB	361	AAATTGTGGCTTCCCGGTTGTCACCGCCAGCAGTGCACGGAGAGAGTTGCTGTTTGA	420
QY	421	TGACAGTGTCCGGGATTCCTGCTGCTTCCACCCCATCGCCATCGAGAACACTCAAGA	480
DB	421	TGACAGTGTCCGGGATTCCTGCTGCTTCCACCCCATCGCCATCGAGAACACTCAAGA	480
QY	481	AGAAAGATGTCCTTTAACTAGTATGATCCGGCTGTCTAAACAAAGCCGAAAGGAGCTGA	540
DB	481	AGAAAGATGTCCTTTAACTAGTATGATCCGGCTGTCTAAACAAAGCCGAAAGGAGCTGA	540
QY	541	GTTGGCTGTGCGACCGCTGAGCAATACTAGCATACCCCTTTGGGCGCTCTAAACGGGT	600
DB	541	GTTGGCTGTGCGACCGCTGAGCAATACTAGCATACCCCTTTGGGCGCTCTAAACGGGT	600
QY	601	CTTGAGGGGTTTTTGTCTGAAAGGAGAACTATATCCGATGACCTGCAGGCAAGCTCTA	660
DB	601	CTTGAGGGGTTTTTGTCTGAAAGGAGAACTATATCCGATGACCTGCAGGCAAGCTCTA	660
QY	661	GAATCGATACGATTTTGAAGTGGCAACAGATATAAATAAAGCAGTTTAAATTTGTGCTGA	720
DB	661	GAATCGATACGATTTTGAAGTGGCAACAGATATAAATAAAGCAGTTTAAATTTGTGCTGA	720
QY	721	ACTTTTAAACAAGCAATCAATTCATGTCGCAACAGATAGCGACAGAGAAAGCGGAAA	780
DB	721	ACTTTTAAACAAGCAATCAATTCATGTCGCAACAGATAGCGACAGAGAAAGCGGAAA	780
QY	781	CAITGGCTGCTCGATCATTCATAAGCAATGCTTTTCTAAAGATTAACGTTATTAAG	840
DB	781	CAITGGCTGCTCGATCATTCATAAGCAATGCTTTTCTAAAGATTAACGTTATTAAG	840
QY	841	ACTATGGATCAATAGTTTGAAGAAAGATGTGATCCGTAGCGGTTTTTCAAAATTTGCAACC	900
DB	841	ACTATGGATCAATAGTTTGAAGAAAGATGTGATCCGTAGCGGTTTTTCAAAATTTGCAACC	900
QY	901	AGGAATGAATTAATCTATCCCTTTTATCAAGAGCGCAAAAGAAAAACGAATGATACACCA	960

Db 1981 GGAAATGATAGATAAAATTTATTAGGTATACCTACTGACAGCTTCCAAGGAGCTAAAGAGGTCCC 2040  
QY 2041 TAGCGCTCTTATCATGCGGAAGCTCGGATCATATGCAAGACAAATAAATTAACCTCGCAACAGC 2100  
Db 2041 TAGCGCTCTTATCATGCGGAAGCTCGGATCATATGCAAGACAAATAAATTAACCTCGCAACAGC 2100  
QY 2101 ACTTGAGAAATGCGACGAATCGAGAAACCCCTCTTTACGCTGGATTAACATATCTAAATAA 2160  
Db 2101 ACTTGAGAAATGCGACGAATCGAGAAACCCCTCTTTACGCTGGATTAACATATCTAAATAA 2160  
QY 2161 AGCGTAAGGAGACGGGTTCAAAAAGGTTTAAATAAAGGAGAGCAATCAATGCAATTAGC 2220  
Db 2161 AGCGTAAGGAGACGGGTTCAAAAAGGTTTAAATAAAGGAGAGCAATCAATGCAATTAGC 2220  
QY 2221 TAGAACTATATTTTGGCAACGCTGGAGAAATTTAGAGAACGCTGCTCTCCAAGACCACTT 2280  
Db 2221 TAGAACTATATTTTGGCAACGCTGGAGAAATTTAGAGAACGCTGCTCTCCAAGACCACTT 2280  
QY 2281 ACAAGAGCTAGTGCACTAAACATTAATTAACGCTATAAGTGTGAGAACACCTGTATA 2340  
Db 2281 ACAAGAGCTAGTGCACTAAACATTAATTAACGCTATAAGTGTGAGAACACCTGTATA 2340  
QY 2341 TAGGAAAGCGGTAGAGAAATTAACCAAGAGAGAGAAATTTAGAGAAATTTAATGCC 2400  
Db 2341 TAGGAAAGCGGTAGAGAAATTAACCAAGAGAGAGAAATTTAGAGAAATTTAATGCC 2400  
QY 2401 ATATGCGTGGCGTTAGGATGGAAACATATCAATTTTCTTGAGAAATCAAAATTTGAAGG 2460  
Db 2401 ATATGCGTGGCGTTAGGATGGAAACATATCAATTTTCTTGAGAAATCAAAATTTGAAGG 2460  
QY 2461 ATTACATGACACTGGCAATGAATTTAGTCTCTTACGTATTAAGAGAGCGGTTTATTC 2520  
Db 2461 ATTACATGACACTGGCAATGAATTTAGTCTCTTACGTATTAAGAGAGCGGTTTATTC 2520  
QY 2521 TTAATATAACGGCTCTTTTATAGAAAAATCCTTAGCGTGGTTTCTTCCGAAATGCTG 2580  
Db 2521 TTAATATAACGGCTCTTTTATAGAAAAATCCTTAGCGTGGTTTCTTCCGAAATGCTG 2580  
QY 2581 GCGGTACCCCAAGATTTAGAAATGAGTAGATCAAAATTTATCTCGAATGAATTCAGGAAAA 2640  
Db 2581 GCGGTACCCCAAGATTTAGAAATGAGTAGATCAAAATTTATCTCGAATGAATTCAGGAAAA 2640  
QY 2641 TCAGATCCCAACATTAACATTAACCAAAATGCAAAATGCAAAATGCAAAATGCAAAATGCAAA 2700  
Db 2641 TCAGATCCCAACATTAACATTAACCAAAATGCAAAATGCAAAATGCAAAATGCAAAATGCAAA 2700  
QY 2701 GTGGATTTAATCCCAATGAGCCAAACAGAACCCAGAGCCAGAACCAAGATCAGAACAAAGTA 2760  
Db 2701 GTGGATTTAATCCCAATGAGCCAAACAGAACCCAGAGCCAGAACCAAGATCAGAACAAAGTA 2760  
QY 2761 ACATTGGATTTAGAAATGGAAG 2812  
Db 2761 ACATTGGATTTAGAAATGGAAG 2812  
QY 2813 TGCAAGAAATCGTTGCTTATTTTAAAGCGGTATACCTAGATATTAACGAAACCAAG 2872  
Db 2813 TGCAAGAAATCGTTGCTTATTTTAAAGCGGTATACCTAGATATTAACGAAACCAAG 2872  
QY 2873 AACTGAATAGAAACGAAAG 2932  
Db 2873 AACTGAATAGAAACGAAAG 2932  
QY 2933 TAAATGATAGCCGATAGATTTGCAACCAACCGCTTATAGTGTAGTGTAGTGTAGTGTAGTGTAG 2992  
Db 2933 TAAATGATAGCCGATAGATTTGCAACCAACCGCTTATAGTGTAGTGTAGTGTAGTGTAGTGTAG 2992  
QY 2993 TCCTCGTAAAGAGTTTATTAATTAATTTGTTGAGAGCGGTATATAACCGTACTATCA 3052  
Db 2993 TCCTCGTAAAGAGTTTATTAATTAATTTGTTGAGAGCGGTATATAACCGTACTATCA 3052  
QY 3053 TTATATAGGAAATCAGAGAGTTTCAAGTATCTAGCTACTGAAATTTAAGATTTGTTAA 3112  
Db 3053 TTATATAGGAAATCAGAGAGTTTCAAGTATCTAGCTACTGAAATTTAAGATTTGTTAA 3112  
QY 3061 TTATATAGGAAATCAGAGAGTTTCAAGTATCTAGCTACTGAAATTTAAGATTTGTTAA 3120  
Db 3061 TTATATAGGAAATCAGAGAGTTTCAAGTATCTAGCTACTGAAATTTAAGATTTGTTAA 3120

QY 3113 GCAATCAATCGGAAATCGTTTGGATGCTTTTGTATTGATTTATATAGAGGTGGAGTTT 3172  
Db 3121 GCAATCAATCGGAAATCGTTTGGATGCTTTTGTATTGATTTATATAGAGGTGGAGTTT 3180  
QY 3173 GTATGAATCATGATGAATGTAAACCTTATATAAAAAATAGTTTATTGGAGATAGAAAAAT 3232  
Db 3181 GTATGAATCATGATGAATGTAAACCTTATATAAAAAATAGTTTATTGGAGATAGAAAAAT 3240  
QY 3233 TAGCAATATCTATACACTAGAAACGTTTAAAGAAAGAGTTTAGAAAAAGAGAAATATCTACT 3292  
Db 3241 TAGCAATATCTATACACTAGAAACGTTTAAAGAAAGAGTTTAGAAAAAGAGAAATATCTACT 3300  
QY 3293 TAGAAAAAATAATCAGATATATTTTCTTCCGAGGGGGAAGATTATATATAAAGTTAA 3352  
Db 3301 TAGAAAAAATAATCAGATATATTTTCTTCCGAGGGGGAAGATTATATATAAAGTTAA 3360  
QY 3353 TAGAAAAAATAATTAATTTTATTCGATTTAGTGGAAAAATTTGACTTATAAAGGAAAAA 3412  
Db 3361 TAGAAAAAATAATTAATTTTATTCGATTTAGTGGAAAAATTTGACTTATAAAGGAAAAA 3420  
QY 3413 AATCTTTTCAAAACATGCAATATTGAACAGATTGAATGAAAAAGCAAAACCAAGTTAAT 3472  
Db 3421 AATCTTTTCAAAACATGCAATATTGAACAGATTGAATGAAAAAGCAAAACCAAGTTAAT 3480  
QY 3473 AAACAACTATTTTATAGGATTTATAGGAAAGAGAGAAACAGCTGAATGAATATCCCTTTT 3532  
Db 3481 AAACAACTATTTTATAGGATTTATAGGAAAGAGAGAAACAGCTGAATGAATATCCCTTTT 3540  
QY 3533 TTGTAGAAAACTGTGCTTATGACGGCTTGTAAAGTACAAATTTAAAAATAGTAAAAATTC 3592  
Db 3541 TTGTAGAAAACTGTGCTTATGACGGCTTGTAAAGTACAAATTTAAAAATAGTAAAAATTC 3600  
QY 3593 GCTCAATCACTACCAAGCCAGGTAAGAAAGAGGGGCTATTTTTGCTATCGCTCAAAAT 3652  
Db 3601 GCTCAATCACTACCAAGCCAGGTAAGAAAGAGGGGCTATTTTTGCTATCGCTCAAAAT 3660  
QY 3653 CAACATGATTTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3712  
Db 3661 CAACATGATTTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3720  
QY 3713 AAGATACATTTTACACATTTGACACCCCAACCGTTTATCGTTATGGAACGATATGCAAGCAAAA 3772  
Db 3721 AAGATACATTTTACACATTTGACACCCCAACCGTTTATCGTTATGGAACGATATGCAAGCAAAA 3780  
QY 3773 ACCGTTTATACACGAAAGAGACATTTGAAACCAATTTAAGACAAATCAATACCTCTTTA 3832  
Db 3781 ACCGTTTATACACGAAAGAGACATTTGAAACCAATTTAAGACAAATCAATACCTCTTTA 3840  
QY 3833 TTGATTTTGTATTTACACGCGCAAAAGAAACTATTTTCAAGCAAGCGATATTTTAAACACCG 3892  
Db 3841 TTGATTTTGTATTTACACGCGCAAAAGAAACTATTTTCAAGCAAGCGATATTTTAAACACCG 3900  
QY 3893 CTATGATTTAGGTTTATGCTTACTATGATTAATCAAAATCTGATAAAGGTTATCAAGCAT 3952  
Db 3901 CTATGATTTAGGTTTATGCTTACTATGATTAATCAAAATCTGATAAAGGTTATCAAGCAT 3960  
QY 3953 ATTTTGTTTTAAAGAACCCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4012  
Db 3961 ATTTTGTTTTAAAGAACCCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4020  
QY 4013 CAGCAAAATAATTTCCGCAAAATATCCGAGAAATTTTGGAAAGTCTTTGCCAGTTGATC 4072  
Db 4021 CAGCAAAATAATTTCCGCAAAATATCCGAGAAATTTTGGAAAGTCTTTGCCAGTTGATC 4080  
QY 4073 TAACTGTATCATTTTGGTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4132  
Db 4081 TAACTGTATCATTTTGGTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4140  
QY 4133 ATCTTAATACCGTTATTTTCAAGAAATGCGCAAGATTTGCTCTTTCAAAACCAACAGATA 4192  
Db 4141 ATCTTAATACCGTTATTTTCAAGAAATGCGCAAGATTTGCTCTTTCAAAACCAACAGATA 4200

QY 4193 ATAAGGGCTTTACTCGTTCAAGTCTAACGGTTTTAAGCGGTACAGAGGCCAAAAACAAG 4252  
DB 4201 ATAAGGGCTTTACTCGTTCAAGTCTAACGGTTTTAAGCGGTACAGAGGCCAAAAACAAG 4260  
QY 4253 TAGATGAACCTCGTTTAACTCTTATTTGACGAAAGAAATTTTCAGGAGAAAGGGTT 4312  
DB 4261 TAGATGAACCTCGTTTAACTCTTATTTGACGAAAGAAATTTTCAGGAGAAAGGGTT 4320  
QY 4313 TAAATAGGGCGTAATAAGCTCATGTTTACCTCTCTTTAGCTACTTTAGTTTCAAGGCTATT 4372  
DB 4321 TAAATAGGGCGTAATAAGCTCATGTTTACCTCTCTTTAGCTACTTTAGTTTCAAGGCTATT 4380  
QY 4373 CAATTCGAAACGTCGGAATATAATGTTTGAAGTTTAAATTCGATTTAGATCAACCTTAG 4432  
DB 4381 CAATTCGAAACGTCGGAATATAATGTTTGAAGTTTAAATTCGATTTAGATCAACCTTAG 4440  
QY 4433 AAGAAAAAGAGTAATCAAAATTTGTTAGAAGTGCCTATTTCAGAAAACTATCAAGGGGCTA 4492  
DB 4441 AAGAAAAAGAGTAATCAAAATTTGTTAGAAGTGCCTATTTCAGAAAACTATCAAGGGGCTA 4500  
QY 4493 ATAGGGAATACATTACCATTTTTCGAAAGCTTTGGGTATCAAGTGATTTTAAACAGTAAAG 4552  
DB 4501 ATAGGGAATACATTACCATTTTTCGAAAGCTTTGGGTATCAAGTGATTTTAAACAGTAAAG 4560  
QY 4553 ATTATTTTTCGCTCAAGGGTGGTTTAAATTCAGAAAAAGAGCAAGCTCAACGTG 4612  
DB 4561 ATTATTTTTCGCTCAAGGGTGGTTTAAATTCAGAAAAAGAGCAAGCTCAACGTG 4620  
QY 4613 TTCAATTTTGTGAGATGGAAGAGATTTAATGCTTATATTAGCGAAAAAGAGCGATGAT 4672  
DB 4621 TTCAATTTTGTGAGATGGAAGAGATTTAATGCTTATATTAGCGAAAAAGAGCGATGAT 4680  
QY 4673 ACAAGCCTTATTAGTCAGACCAAAAGAGATTTAGAGAGTGTCTAGGCAATTCCTGAAC 4732  
DB 4681 ACAAGCCTTATTAGTCAGACCAAAAGAGATTTAGAGAGTGTCTAGGCAATTCCTGAAC 4740  
QY 4733 GGACATTTAGATAAATTCCTGAAGGTAAGAGCGAATCAGGAAATTTCTTTAAGATTA 4792  
DB 4741 GGACATTTAGATAAATTCCTGAAGGTAAGAGCGAATCAGGAAATTTCTTTAAGATTA 4800  
QY 4793 AACAGGAAGAAATGTTGGCAATTCACCTGCTAGTGTATAATCATTTGTTGCTATCGATCA 4852  
DB 4801 AACAGGAAGAAATGTTGGCAATTCACCTGCTAGTGTATAATCATTTGTTGCTATCGATCA 4860  
QY 4853 TTAAGTAAAGAAAGAAAGAAAGAAAGCTATATAAGGGCGTACAAAATTCCTTTGACT 4912  
DB 4861 TTAAGTAAAGAAAGAAAGAAAGAAAGCTATATAAGGGCGTACAAAATTCCTTTGACT 4920  
QY 4913 TAGAGCATACATTCATTCAGAGACTTTAAACAAGCTTAGCAGAACGCCCTTAAACCGGACA 4972  
DB 4921 TAGAGCATACATTCATTCAGAGACTTTAAACAAGCTTAGCAGAACGCCCTTAAACCGGACA 4980  
QY 4973 CACAACTCGATTTGTTTAGCTATGATACAGGCTGAAATTAACCCGCACTATGCCATTA 5032  
DB 4981 CACAACTCGATTTGTTTAGCTATGATACAGGCTGAAATTAACCCGCACTATGCCATTA 5040  
QY 5033 CATTTATATCTATGATACGTTGTTTGTCTTTTCTTTGCTGTTTACGGAATGATTTAGCAGAA 5092  
DB 5041 CATTTATATCTATGATACGTTGTTTGTCTTTTCTTTGCTGTTTACGGAATGATTTAGCAGAA 5100  
QY 5093 ATATACAGAGTACATTTTAAATTAATTTAGGGGAGAGAGAGTAGGCCGGAAC 5152  
DB 5101 ATATACAGAGTACATTTTAAATTAATTTAGGGGAGAGAGAGTAGGCCGGAAC 5160  
QY 5153 TTTTAGTTGCTTGGACTGAACGAGTGTAGGGAAGGCTACTAAAACGTCGAGGGGCACT 5212  
DB 5161 TTTTAGTTGCTTGGACTGAACGAGTGTAGGGAAGGCTACTAAAACGTCGAGGGGCACT 5220  
QY 5213 GAGAGCGAAGCAACCTTGAATTTTAAATTTCTATCTTTTATAGGTCATTTAGATATA 5272  
DB 5221 GAGAGCGAAGCAACCTTGAATTTTAAATTTCTATCTTTTATAGGTCATTTAGATATA 5280  
QY 5273 CTTATTTGCTCTATAAATCTTTTAGCAGCATATAAGATTTATTTGAATAGTGTCAATTAAGT 5332

DB 5281 CTTATTTGCTCTATAAATCTATTAGCAGCATATAAGATTTATTGAAATAGTCAATTTAAGT 5340  
QY 5333 TCAGCATATTAGAGGAGGAAATCTTCGAGAAATATTGAGAGACCCGATTTACATGGAAT 5392  
DB 5341 TCAGCATATTAGAGGAGGAAATCTTCGAGAAATATTGAGAGACCCGATTTACATGGAAT 5400  
QY 5393 CGATTAGTCTTGTGTTTACGTTGTTTAACTAAAAAGTAGTGAATTTTGAATTTTGGT 5452  
DB 5401 CGATTAGTCTTGTGTTTACGTTGTTTAACTAAAAAGTAGTGAATTTTGAATTTTGGT 5460  
QY 5453 GTGTGTGCTTGTGTTTGTAGTATTTCTAGTCAAAAGTGAATTAATA 5497  
DB 5461 GTGTGTGCTTGTGTTTGTAGTATTTCTAGTCAAAAGTGAATTAATA 5505

RESULT 2  
AR381898  
LOCUS AR381898 5217 bp DNA linear PAT 18-DEC-2003  
DEFINITION Sequence 25 from patent US 6610300.  
ACCESSION AR381898  
VERSION AR381898.1 GI:40090247  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 5217)  
AUTHORS Segers, R.P.A.M., Waterfield, N.R., Frandsen, P.L. and Wells, J.M.  
TITLE Clostridium perfringens vaccine  
JOURNAL Patent: US 6610300-A 25 AUG-2003;  
FEATURES  
Location/Qualifiers  
1..5217  
source  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 90.9%; Score 4994.4; DB 6; Length 5217;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 4998; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 494 TTCTAACTAGTAGATCCGGCTGCTAACAAAGCCCGAAGAGAGCTGAGTTGGCTGCTGCC 553  
DB 214 TCCAAATGATCAGATCCGGCTGCTAACAAAGCCCGAAGAGAGCTGAGTTGGCTGCTGCC 273  
QY 554 ACCGCTGAGCAATAACTAGCATTAACCCCTTGGGGCCCTCTAAACGGGCTCTTGAAGGGGTTTT 613  
DB 274 ACCGCTGAGCAATAACTAGCATTAACCCCTTGGGGCCCTCTAAACGGGCTCTTGAAGGGGTTTT 333  
QY 614 TTGCTGAAAGAGGAGAACTATATCCGGATGACCTGCGAGCAAGCTCTAGAAATCGATACGAT 673  
DB 334 TTGCTGAAAGAGGAGAACTATATCCGGATGACCTGCGAGCAAGCTCTAGAAATCGATACGAT 393  
QY 674 TTGAGTGGCAACAGATATAAAGAGCAGTTTAAATTTGTTGCTGAACCTTTTAAACAA 733  
DB 394 TTGAGTGGCAACAGATATAAAGAGCAGTTTAAATTTGTTGCTGAACCTTTTAAACAA 453  
QY 734 GCAAATCAATTCATTGTCGCAACAGATAGCGACAGAGAGCGGCAAAACATTTGCTGCTG 793  
DB 454 GCAAATCAATTCATTGTCGCAACAGATAGCGACAGAGAGCGGCAAAACATTTGCTGCTG 513  
QY 794 ATCAATTCATAAGCAATGCTTTTCTAAGATTAACAGCTATAAAGACTATGATCAAT 853  
DB 514 ATCAATTCATAAGCAATGCTTTTCTAAGATTAACAGCTATAAAGACTATGATCAAT 573  
QY 854 AGTTTAGAAAAAGATGTGATCCGCTAGCGGTTTTTCAAAATTTGCAACAGCAATTAATAC 933  
DB 574 AGTTTAGAAAAAGATGTGATCCGCTAGCGGTTTTTCAAAATTTGCAACAGCAATTAATAC 633  
QY 914 TATCCCTTTTATCAAGAGCGCAAAAGAAAAAGAAATGATACCAATCAGTCAGTCAAAA 973  
DB 634 TATCCCTTTTATCAAGAGCGCAAAAGAAAAAGAAATGATACCAATCAGTCAGTCAAAA 693  
QY 974 AAGATATAATGGGAGATAAGACGGTTTCTGTTGCTGACTTTCGACCATATCATATAAAA 1033



Db 694 AAGATATAATGGGAGATAAGACGGTTCGTCTCGTCTGACCTTGCACCATATCATATAAA 753  
Qy 1034 TCGAARACAGCAAGAAAGTGGCGGAAACGTAAAGAGTTTAAAGAAATAGAGCTTAGAGCA 1093  
Db 754 TCGAARACAGCAAGAAAGTGGCGGAAACGTAAAGAGTTTAAAGAAATAGAGCTTAGAGCA 813  
Qy 1094 AACTTAAGAGTGTGTGATAGTGCAGTATCTTAAATTTTGTATATAAGAAATGAAGTT 1153  
Db 814 AACTTAAGAGTGTGTGATAGTGCAGTATCTTAAATTTTGTATATAAGAAATGAAGTT 873  
Qy 1154 AATTAAGATCTTAAATTTTAAATTTAAAGAGGAGTATACATGAACAAATAATATAAA 1213  
Db 874 AATTAAGATCTTAAATTTTAAATTTAAAGAGGAGTATACATGAACAAATAATATAAA 933  
Qy 1214 TATTTCTCAAACTTTTAAAGAGTGAAGAAAGTACTCAACCAATAATAAAACAATGAAT 1273  
Db 934 TATTTCTCAAACTTTTAAAGAGTGAAGAAAGTACTCAACCAATAATAAAACAATGAAT 993  
Qy 1274 TTAAGAGAAACCGATACCGTTTACGAAATTTGGAACAGGTAAGGGCAATTAACAGCAAA 1333  
Db 994 TTAAGAGAAACCGATACCGTTTACGAAATTTGGAACAGGTAAGGGCAATTAACAGCAAA 1053  
Qy 1334 CTGCTTAAATTAAGTAAACAGGTAACGTCTATTGAAATTAAGACAGTCACTATTCTCACTTA 1393  
Db 1054 CTGCTTAAATTAAGTAAACAGGTAACGTCTATTGAAATTAAGACAGTCACTATTCTCACTTA 1113  
Qy 1394 TCGTCAAGAAATTTAAACACTGAATCTCGTGTCACTTTAATTCACCAAGATATTTCTACAG 1453  
Db 1114 TCGTCAAGAAATTTAAACACTGAATCTCGTGTCACTTTAATTCACCAAGATATTTCTACAG 1173  
Qy 1454 TTTCAATTCCTCAACCAACAGGTAATAAATTTGGGAGTATTCCTTACCANTTTAAGC 1513  
Db 1174 TTTCAATTCCTCAACCAACAGGTAATAAATTTGGGAGTATTCCTTACCANTTTAAGC 1233  
Qy 1514 ACACAAATTTAAAGAGTGTGTGTGAAAGCCATCGCTCTGACATCTATCTGATTTT 1573  
Db 1234 ACACAAATTTAAAGAGTGTGTGTGAAAGCCATCGCTCTGACATCTATCTGATTTT 1293  
Qy 1574 GAAGAGGATTTCAAGAGTACTTGGATATTTCAACCAAGATATTTCTACAG 1633  
Db 1294 GAAGAGGATTTCAAGAGTACTTGGATATTTCAACCAAGATATTTCTACAG 1353  
Qy 1634 ACTCAAGTCTCGATTCAGCAATTTGCTTAAGCTCCAGCGGAATGCTTTCATCTCAAAACCA 1693  
Db 1354 ACTCAAGTCTCGATTCAGCAATTTGCTTAAGCTCCAGCGGAATGCTTTCATCTCAAAACCA 1413  
Qy 1694 AAAGTAAACAGTCTCTTAATAAACTTACCCGCCATACCACAGATGTTCCAGATAAATAT 1753  
Db 1414 AAAGTAAACAGTCTCTTAATAAACTTACCCGCCATACCACAGATGTTCCAGATAAATAT 1473  
Qy 1754 TGGAGCTATATACGTACTTGTGTTTCAAAATGGTCAATCGGAATATCGTCAACGTGTT 1813  
Db 1474 TGGAGCTATATACGTACTTGTGTTTCAAAATGGTCAATCGGAATATCGTCAACGTGTT 1533  
Qy 1814 ACTAAAAATCAGTTTTCATCAAGCAATGAACACGCCAAAGTAAACAAATTTAAAGTACCGTT 1873  
Db 1534 ACTAAAAATCAGTTTTCATCAAGCAATGAACACGCCAAAGTAAACAAATTTAAAGTACCGTT 1593  
Qy 1874 ACTTATGAGCAAGTATGCTATTTTAAATAGTATCTATTAATTTAAACGGGAGAAATAA 1933  
Db 1594 ACTTATGAGCAAGTATGCTATTTTAAATAGTATCTATTAATTTAAACGGGAGAAATAA 1653  
Qy 1934 TTCTATGAGTGGTGTGTTTAAATTTGGAAGTTTACACGTTTACTAAAGGGAATGTAGATAA 1993  
Db 1654 TTCTATGAGTGGTGTGTTTAAATTTGGAAGTTTACACGTTTACTAAAGGGAATGTAGATAA 1713  
Qy 1994 ATTATAGGTATATCTAGACGTTTCCAAAGGAGTAAAGAGTCCCTAGCGCTCTTATC 2053  
Db 1714 ATTATAGGTATATCTAGACGTTTCCAAAGGAGTAAAGAGTCCCTAGCGCTCTTATC 1773  
Qy 2054 ATCGGAAGCTCGGATCATATGCAAGCAAAATPAACTCGCAACAGCACTTCGGAAGATG 2113

Db 1774 ATGGGAAAGCTCGGATCATATGCAAGACAAAATAAATCGCAACAGCACTTTGGAGAAATG 1833  
Qy 2114 GGACCAATTCGAGAAACCCCTCTTACGCTGANTTACATATCTAATAAAGCCGTAAAGGAGA 2173  
Db 1834 GGACCAATTCGAGAAACCCCTCTTACGCTGANTTACATATCTAATAAAGCCGTAAAGGAGA 1893  
Qy 2174 CGGTTTCAAAAGGTTTAAATTAAGGAGAGCAATCAATGCAATTTAGCTAGAACTATATTT 2233  
Db 1894 CGGTTTCAAAAGGTTTAAATTAAGGAGAGCAATCAATGCAATTTAGCTAGAACTATATTT 1953  
Qy 2234 TTTGGACACCGTGGAGAAATTTAGAGAACGCTGCTCCGAAGACCACTTCAAGAGCTAGT 2293  
Db 1954 TTTGGACACCGTGGAGAAATTTAGAGAACGCTGCTCCGAAGACCACTTCAAGAGCTAGT 2013  
Qy 2294 GCACCTAAACATAATTTATTAAACGCTATAGTGTGGAAACGCTGATATATATGAAAGAGCC 2353  
Db 2014 GCACCTAAACATAATTTATTAAACGCTATAGTGTGGAAACGCTGATATATATGAAAGAGCC 2073  
Qy 2354 GTAGAGAAATTTAAAGCAAGAGAGAGAAATTTAGAGAGAAATTTAATGCCATATGCGTGGCG 2413  
Db 2074 GTAGAGAAATTTAAAGCAAGAGAGAGAAATTTAGAGAGAAATTTAATGCCATATGCGTGGCG 2133  
Qy 2414 TTAGGATCGGACATATCAATTTCTTGGAGAAATCAAAATTTGAAGGATTAACATGACACT 2473  
Db 2134 TTAGGATCGGACATATCAATTTCTTGGAGAAATCAAAATTTGAAGGATTAACATGACACT 2193  
Qy 2474 GGGCAAAATTTAAGCTCTTTACGTATAAAGAGCGGTTTTATTCTTAATAACGGC 2533  
Db 2194 GGGCAAAATTTAAGCTCTTTACGTATAAAGAGCGGTTTTATTCTTAATAACGGC 2253  
Qy 2534 TCTTTTATAGAAATAATTCCTAGCTGTTTTTTCGGAATGCTGCGGTACCCCAAG 2593  
Db 2254 TCTTTTATAGAAATAATTCCTAGCTGTTTTTTCGGAATGCTGCGGTACCCCAAG 2313  
Qy 2594 AATTAGAAATGAGTAGATCAAAATTTATTCAGAAATAGAAATCAGGAAATCAGATCAACCA 2653  
Db 2314 AATTAGAAATGAGTAGATCAAAATTTATTCAGAAATAGAAATCAGGAAATCAGATCAACCA 2373  
Qy 2654 TAAAAAACHACTAGAACAAATTTGCAAGTTAATCTAACTCAACCTAGTAGTGAATTTATCC 2713  
Db 2374 TAAAAAACHACTAGAACAAATTTGCAAGTTAATCTAACTCAACCTAGTAGTGAATTTATCC 2433  
Qy 2714 CAATGAGCCAAACAGAACAGAGCCAGAAACAGAAATCAGAACTAGAACTTGAATTTAG 2773  
Db 2434 CAATGAGCCAAACAGAACAGAGCCAGAAACAGAAATCAGAACTAGAACTTGAATTTAG 2493  
Qy 2774 AAATGGAAGAGAAAGAAAGCAATGACTTCTGTTGAATTAATGACGAAATCGTTCCTTAT 2833  
Db 2494 AAATGGAAGAGAAAGAAAGCAATGACTTCTGTTGAATTAATGACGAAATCGTTCCTTAT 2553  
Qy 2834 TTTTAAAAAGCGGTATACATAGATATAACGAAACAAAGCACTGAATAGAAACGAAAGAA 2893  
Db 2554 TTTTAAAAAGCGGTATACATAGATATAACGAAACAAAGCACTGAATAGAAACGAAAGAA 2613  
Qy 2894 GAGCCATGACACATTTTAAATGTTTGAAGCAATTTTGAAGCAATTTTAAATGCAATAGCA 2953  
Db 2614 GAGCCATGACACATTTTAAATGTTTGAAGCAATTTTGAAGCAATTTTAAATGCAATAGCA 2673  
Qy 2954 TTGCCAAACCAACCGTTATCAGTTAGTCAAGTGAACCTTCCCTCGTGAAGAGTTATTTA 3013  
Db 2674 TTGCCAAACCAACCGTTATCAGTTAGTCAAGTGAACCTTCCCTCGTGAAGAGTTATTTA 2733  
Qy 3014 ATTAACCTTTGTTTGAAGAGCGGTATATAACCGTACTACTCATTTATATAGGGAATCAGAG 3073  
Db 2734 ATTAACCTTTGTTTGAAGAGCGGTATATAACCGTACTACTCATTTATATAGGGAATCAGAG 2793  
Qy 3074 TTTTCAAGTATCTAAGCTACTGAATTTTAAAGATTTGTTAAGCAATCAATCGGAAATCGTTT 3133  
Db 2794 TTTTCAAGTATCTAAGCTACTGAATTTTAAAGATTTGTTAAGCAATCAATCGGAAATCGTTT 2853  
Qy 3134 GATTGCTTTTTTCTGTTTCTATTATAGAGTGGAGTTTGTATGAATCATGATGAATGTA 3193  
Db 2854 GATTGCTTTTTTCTGTTTCTATTATAGAGTGGAGTTTGTATGAATCATGATGAATGTA 2913

QY 3194 AAACTTATATAAAATAGTTTATTGAGATAGAAAATTTAGCAATATCTATACACTAG 3253  
DB 2914 AAACTTATATAAAATAGTTTATTGAGATAGAAAATTTAGCAATATCTATACACTAG 2973  
QY 3254 AAACGTTTAAAGAAAGAGTTAGAAAGAGAAATATCTACTTTAGAAAACAAATCAGATAAGT 3313  
DB 2974 AAACGTTTAAAGAAAGAGTTAGAAAGAGAAATATCTACTTTAGAAAACAAATCAGATAAGT 3033  
QY 3314 ATTTTCTTCGAGGGGAGATATATATATATAGTTTAAATAGAAAATAACAAATATTTT 3373  
DB 3034 ATTTTCTTCGAGGGGAGATATATATATATAGTTTAAATAGAAAATAACAAATATTTT 3093  
QY 3374 ATTCGATTAGTGGAAAATAATTCGATTATATAAGGAAAATAATCTTTTCAAAAACATGCAA 3433  
DB 3094 ATTCGATTAGTGGAAAATAATTCGATTATATAAGGAAAATAATCTTTTCAAAAACATGCAA 3153  
QY 3434 TATTGAAACAGTTGAAATGAAAGCAACCAAGTTAAATTAATAACCACTATTTATAGGAT 3493  
DB 3154 TATTGAAACAGTTGAAATGAAAGCAACCAAGTTAAATTAATAACCACTATTTATAGGAT 3213  
QY 3494 TTATAGGAAAGGAGAAACAGCTGAATGAATCCCTTTTGTGTAGAACTGTGCTTCATG 3553  
DB 3214 TTATAGGAAAGGAGAAACAGCTGAATGAATCCCTTTTGTGTAGAACTGTGCTTCATG 3273  
QY 3554 ACGCTTGTGTTAAAGTACAAATTTTAAATAATAGTAAATTCGCTCAATCACTACCAAGCCAG 3613  
DB 3274 ACGCTTGTGTTAAAGTACAAATTTTAAATAATAGTAAATTCGCTCAATCACTACCAAGCCAG 3333  
QY 3614 GTAAAGCAAGGGCTATTTTTCGCTATCGCTCAAAATCAAGCATGATGGCGCTCGT 3673  
DB 3334 GTAAAGCAAGGGCTATTTTTCGCTATCGCTCAAAATCAAGCATGATGGCGCTCGT 3393  
QY 3674 GTGTGTTCTGACTTCGAGGAAGCGATTCAAGAAAATCAAGATACATTTTACATGGA 3733  
DB 3394 GTGTGTTCTGACTTCGAGGAAGCGATTCAAGAAAATCAAGATACATTTTACATGGA 3453  
QY 3734 CACCAAGCTTATCGTTATGGAAGCTATGAGCAAGAAACCGTTATACACCAAGGAC 3793  
DB 3454 CACCAAGCTTATCGTTATGGAAGCTATGAGCAAGAAACCGTTATACACCAAGGAC 3513  
QY 3794 ATTCGAAAACAAATTAAGCAAAATCAATACCTCTTTTATGATTTTATGATTTTACACCG 3853  
DB 3514 ATTCGAAAACAAATTAAGCAAAATCAATACCTCTTTTATGATTTTATGATTTTACACCG 3573  
QY 3854 CAAAGAAACTATTTTCAGCAAGCGATATTTTAAACACCGCTATTTGATTTTATGCTTATGC 3913  
DB 3574 CAAAGAAACTATTTTCAGCAAGCGATATTTTAAACACCGCTATTTGATTTTATGCTTATGC 3633  
QY 3914 CTACTATGATTTATCAATCTCATAAAGTTATCAAGCATATTTTGTATTAAGAAACCGCAG 3973  
DB 3634 CTACTATGATTTATCAATCTCATAAAGTTATCAAGCATATTTTGTATTAAGAAACCGCAG 3693  
QY 3974 TCTATGACTTCAAAATCAGAAATTAATCTGTCAAAGCAGCAAAATTAATTCGAAA 4033  
DB 3694 TCTATGACTTCAAAATCAGAAATTAATCTGTCAAAGCAGCAAAATTAATTCGAAA 3753  
QY 4034 ATATCCGAGATATTTTGGAAAGCTTTTCCAGTTGATTAACGTGTAATCATTTTGGTA 4093  
DB 3754 ATATCCGAGATATTTTGGAAAGCTTTTCCAGTTGATTAACGTGTAATCATTTTGGTA 3813  
QY 4094 TTGCTCGCATACCAAGACCGCAATGTAGAAATTTTGTATTCCTPAATTAACGTTATCTT 4153  
DB 3814 TTGCTCGCATACCAAGACCGCAATGTAGAAATTTTGTATTCCTPAATTAACGTTATCTT 3873  
QY 4154 TCAAGATATGCAAGATTTGCTTTTCAACCAAGATTAATAGGGCTTTTACTCGTTCAA 4213  
DB 3874 TCAAGATATGCAAGATTTGCTTTTCAACCAAGATTAATAGGGCTTTTACTCGTTCAA 3933  
QY 4214 GTCTAACCGGTTTAAAGCGGTAACAGAGGCAAAACCAAGTAGATGAACCCCTGGTTTAAATC 4273  
DB 3934 GTCTAACCGGTTTAAAGCGGTAACAGAGGCAAAACCAAGTAGATGAACCCCTGGTTTAAATC 3993

QY 4274 TCTTATTGCAAGAAACGAAATTTTTCAGGAGAAAAGGGTTTAAATAGGCGGTAATAACGTCA 4333  
DB 3994 TCTTATTGCAAGAAACGAAATTTTTCAGGAGAAAAGGGTTTAAATAGGCGGTAATAACGTCA 4053  
QY 4334 TGTTTACCTCTCTTTAGCCTACTTTAGTTCAGGCTATTCAATCGAAACGTGCGAATATA 4393  
DB 4054 TGTTTACCTCTCTTTAGCCTACTTTAGTTCAGGCTATTCAATCGAAACGTGCGAATATA 4113  
QY 4394 ATATGTTTGAAGTTTAAATATCGATTAGATCAACCTTTAGAGAAAAGAAAGTAAATCAAAA 4453  
DB 4114 ATATGTTTGAAGTTTAAATATCGATTAGATCAACCTTTAGAGAAAAGAAAGTAAATCAAAA 4173  
QY 4454 TTGTAGAAAGTGCCTATTTCAGAAAACCTATCAAGGGGCTAATAGGAAATACATTTACCTTC 4513  
DB 4174 TTGTAGAAAGTGCCTATTTCAGAAAACCTATCAAGGGGCTAATAGGAAATACATTTACCTTC 4233  
QY 4514 TTTCGAAAAGCTTGGGTATCAAGTGAATTTAAACGATTAAGATTTTATTTGTCGCTCAAGGGT 4573  
DB 4234 TTTCGAAAAGCTTGGGTATCAAGTGAATTTAAACGATTAAGATTTTATTTGTCGCTCAAGGGT 4293  
QY 4574 GGTTTAAATTTCAAGAAAAGAAAGCGAAACGTCAACGTTTCATTTGTCAGAAATGGAAG 4633  
DB 4294 GGTTTAAATTTCAAGAAAAGAAAGCGAAACGTCAACGTTTCATTTGTCAGAAATGGAAG 4353  
QY 4634 AAGATTTAATGCTTTATATTAGCGAAAAGAAAGCGATGTATACAAGCTTTATTTAGTGACGA 4693  
DB 4354 AAGATTTAATGCTTTATATTAGCGAAAAGAAAGCGATGTATACAAGCTTTATTTAGTGACGA 4413  
QY 4694 CCAAAAAGAGATTAGAGAAGTCTGCTAGCAATTCCTGAACGACATTAGATAAATTTGCTGA 4753  
DB 4414 CCAAAAAGAGATTAGAGAAGTCTGCTAGCAATTCCTGAACGACATTAGATAAATTTGCTGA 4473  
QY 4754 AGGTACTCAAGGCGAATCAGGAAATTTCTTTAGATTAAACAGGAGAAATCGTGCA 4813  
DB 4474 AGGTACTCAAGGCGAATCAGGAAATTTCTTTAGATTAAACAGGAGAAATCGTGCA 4533  
QY 4814 TTCAACTTGTAGTTTAAATCATTGTTGCTATCGATCAATTTAAAGTAAAGAAAGAA 4873  
DB 4534 TTCAACTTGTAGTTTAAATCATTGTTGCTATCGATCAATTTAAAGTAAAGAAAGAA 4593  
QY 4874 AAGAAAGCTATATAAGGCGCTGCAAAATTTCTTTGACTTAGAGCATACATTTCAAG 4933  
DB 4594 AAGAAAGCTATATAAGGCGCTGCAAAATTTCTTTGACTTAGAGCATACATTTCAAG 4653  
QY 4934 AGACTTTTAAACAGCTAGCAGAACGCCCTTAAACAGGACACACAACTCGATTGTTTAGCT 4993  
DB 4654 AGACTTTTAAACAGCTAGCAGAACGCCCTTAAACAGGACACACAACTCGATTGTTTAGCT 4713  
QY 4994 ATGATACAGGCTGAAAATAAACCAGCACTATGCAATTTATCTATGATACGTG 5053  
DB 4714 ATGATACAGGCTGAAAATAAACCAGCACTATGCAATTTATCTATGATACGTG 4773  
QY 5054 TTTGTTTTTTCTTTCTGCTTTAGCGAATGATTAGCAGAAAATATACAGAGTAAGATTTTAA 5113  
DB 4774 TTTGTTTTTTCTTTCTGCTTTAGCGAATGATTAGCAGAAAATATACAGAGTAAGATTTTAA 4833  
QY 5114 TTAATTTATAGGGGAGAGAGAGTAGGCCCGAAAACCTTTTATGTTGGCTTGACCTGAA 5173  
DB 4834 TTAATTTATAGGGGAGAGAGAGTAGGCCCGAAAACCTTTTATGTTGGCTTGACCTGAA 4893  
QY 5174 CGAAGTAGGGAAGAGGCTTACTTAAACGTCGAGGGCAGTGAGAGCAAGCAACCTTGA 5233  
DB 4894 CGAAGTAGGGAAGAGGCTTACTTAAACGTCGAGGGCAGTGAGAGCAAGCAACCTTGA 4953  
QY 5234 TTTTAAATTTTCTATCTTTTATAGGTCATTAGATATCTATTTGTCCTATTAACCTAT 5293  
DB 4954 TTTTAAATTTTCTATCTTTTATAGGTCATTAGATATCTATTTGTCCTATTAACCTAT 5013  
QY 5294 TTAGCAGCATATAGATTTTATTTAGTTTCAATTTAGTTTCAATTTAGGAGGAGAA 5353  
DB 5014 TTAGCAGCATATAGATTTTATTTGAATAGGTCATTTTGAATTTAGGAGGAGGAA 5073  
QY 5354 ATCTTGGAGAAATATTTGAAGAACCCGATTACGATTGATTGTTCTTTGTTGGTTACG 5413



QY 2114 GGACGAATCGAGAAACCCCTCTTTACGCTGGATTACATATCTAATAAAGCCGTAAGGAGA 2173  
DB |||||  
QY 1834 GGACGAATCGAGAAACCCCTCTTTACGCTGGATTACATATCTAATAAAGCCGTAAGGAGA 1893  
DB |||||  
QY 2174 CGGTTTCAAAAAGGTTTAAATAAAGGAGAGCAATCAATCAATAGCTAGACTATATTT 2233  
DB |||||  
QY 1894 CGGTTTCAAAAAGGTTTAAATAAAGGAGAGCAATCAATCAATAGCTAGACTATATTT 1953  
DB |||||  
QY 2234 TTTGGACAGCTCGAGAAATTTAGAGACGCTCTCCACGACAGTTTACAAAGAGCTAGT 2293  
DB |||||  
QY 1954 TTTGGACAGCTCGAGAAATTTAGAGACGCTCTCCACGACAGTTTACAAAGAGCTAGT 2013  
DB |||||  
QY 2294 GCATTAACATATTTATTAAGCTTATAGTGTGTGGAACTCTGTATATATGGAAGGCC 2353  
DB |||||  
QY 2014 GCATTAACATATTTATTAAGCTTATAGTGTGTGGAACTCTGTATATATGGAAGGCC 2073  
DB |||||  
QY 2354 GTAGAAGATTTAAAGCAAGAGAGAAATTTAGAGAGATTTAATGCCATATGCTGGCCG 2413  
DB |||||  
QY 2074 GTAGAAGATTTAAAGCAAGAGAGAAATTTAGAGAGATTTAATGCCATATGCTGGCCG 2133  
DB |||||  
QY 2414 TTAGATGGGAACTATCAATTTTCTTTGGAGAAATACAAATTTGAAGGATTTACATGACACT 2473  
DB |||||  
QY 2134 TTAGATGGGAACTATCAATTTTCTTTGGAGAAATACAAATTTGAAGGATTTACATGACACT 2193  
DB |||||  
QY 2474 GGGCAATGAATTTACGTCCTTTACGTTATTAAGAGCGGTTTTATTTCTTAATTAAGCGC 2533  
DB |||||  
QY 2194 GGGCAATGAATTTACGTCCTTTACGTTATTAAGAGCGGTTTTATTTCTTAATTAAGCGC 2253  
DB |||||  
QY 2534 TCTTTTATAGAAAATCCTTAGCTGGTTTTTTTCCGAAATGCTGGCGGTACCCCAAG 2593  
DB |||||  
QY 2254 TCTTTTATAGAAAATCCTTAGCTGGTTTTTTTCCGAAATGCTGGCGGTACCCCAAG 2313  
DB |||||  
QY 2594 AATTAGAAATGAGTAGATCAATTTATTACGAATPAGAAATCAGGAAATCAGATCCAAACCA 2653  
DB |||||  
QY 2314 AATTAGAAATGAGTAGATCAATTTATTACGAATPAGAAATCAGGAAATCAGATCCAAACCA 2373  
DB |||||  
QY 2654 TAAAACTAGAACAAATTTCCAAAGTTTAACTCAAGCTAGTAGTGGATTTAATCC 2713  
DB |||||  
QY 2374 TAAAACTAGAACAAATTTCCAAAGTTTAACTCAAGCTAGTAGTGGATTTAATCC 2433  
DB |||||  
QY 2714 CAATAGACCAACAGAACCCAGAGCCAGAAACAGAAATCAGAACAGTAACATTTGGATTTAG 2773  
DB |||||  
QY 2434 CAATAGACCAACAGAACCCAGAGCCAGAAACAGAAATCAGAACAGTAACATTTGGATTTAG 2493  
DB |||||  
QY 2774 AAATGGAGAGAAAGAAAGCAATGACTTCTGTGGAATTAATGCAGAAATCGTTGCTTATT 2833  
DB |||||  
QY 2494 AAATGGAGAGAAAGAAAGCAATGACTTCTGTGGAATTAATGCAGAAATCGTTGCTTATT 2553  
DB |||||  
QY 2834 TTTTTTAAAGCGGTATCTAGATATAACGAAACCAAGCAATCAATAGAGAAACGAAAAA 2893  
DB |||||  
QY 2554 TTTTTTAAAGCGGTATCTAGATATAACGAAACCAAGCAATCAATAGAGAAACGAAAAA 2613  
DB |||||  
QY 2894 GAGCAATGACACATTTATTAAGATTTTGAAGCAATTTTAAATGCAATGACCGGATTAAGA 2953  
DB |||||  
QY 2614 GAGCAATGACACATTTATTAAGATTTTGAAGCAATTTTAAATGCAATGACCGGATTAAGA 2673  
DB |||||  
QY 2954 TTGCCAACCAACGCTTATCAGTTAGTACAGATGACTTCTCCCTCGTAAGAGATTTATTA 3013  
DB |||||  
QY 2674 TTGCCAACCAACGCTTATCAGTTAGTACAGATGACTTCTCCCTCGTACAGAGTTATTA 2733  
DB |||||  
QY 3014 ATTAACCTTTGTTGAAGCGGTATATAACCGTACTATCATTTATATAGGAAATCAGAGAG 3073  
DB |||||  
QY 2734 ATTAACCTTTGTTGAAGCGGTATATAACCGTACTATCATTTATATAGGAAATCAGAGAG 2793  
DB |||||  
QY 3074 TTTTCAAGTATCTAAGCTACTGAATTTAAGATTTTAAAGATTTTAAAGCAATCAATCGGAATCGTTT 3133  
DB |||||  
QY 2794 TTTTCAAGTATCTAAGCTACTGAATTTAAGATTTTAAAGATTTTAAAGCAATCAATCGGAATCGTTT 2853  
DB |||||  
QY 3134 GATTGCTTTTTTGTGATTTCAATTTATAGAAAGGTGGAGTTTGTGATGAATCATGATGAATGTA 3193  
DB |||||  
QY 2854 GATTGCTTTTTTGTGATTTCAATTTATAGAAAGGTGGAGTTTGTGATGAATCATGATGAATGTA 2913  
DB |||||  
QY 3194 AAACCTTATATAAAAATAGTTTATTTGGAGATAAGAAAATTTAGCAAAATCTATACACTAG 3253

DB 2914 AAACCTTATATAAAAATAGTTTATTTGGAGATAAGAAAATTTAGCAAAATCTATACACTAG 2973  
QY |||||  
DB 3254 AAACCTTATATAAAAATAGTTTATTTGGAGATAAGAAAATTTAGCAAAATCTATACACTAG 3313  
QY |||||  
DB 2974 AAACCTTATATAAAAATAGTTTATTTGGAGATAAGAAAATTTAGCAAAATCTATACACTAG 3033  
QY |||||  
DB 3314 ATTTTCTTTCGAGGAGGAGATTTATATATAAGTTTAAATAGAAAATAACAAAATAATTT 3373  
QY |||||  
DB 3034 ATTTTCTTTCGAGGAGGAGATTTATATATAAGTTTAAATAGAAAATAACAAAATAATTT 3093  
QY |||||  
DB 3374 ATTCGATTTAGTGGAAAAAATTTGACTTATATAAGGAAAAAATCTTTTCAAAACATCAA 3433  
QY |||||  
DB 3094 ATTCGATTTAGTGGAAAAAATTTGACTTATATAAGGAAAAAATCTTTTCAAAACATCAA 3153  
QY |||||  
DB 3434 TATTGAAACAGTTGAATGAAAAAGCAAAACCAAGTTAAATTAACCAACCTATTTTATAGGAT 3493  
DB |||||  
DB 3154 TATTGAAACAGTTGAATGAAAAAGCAAAACCAAGTTAAATTAACCAACCTATTTTATAGGAT 3213  
QY |||||  
DB 3494 TTATAGGAAAGAGAAACAGCTGAATGAATATCCCTTTTGTGTAGAAAACCTGTGCTTCATG 3553  
DB |||||  
DB 3214 TTATAGGAAAGAGAAACAGCTGAATGAATATCCCTTTTGTGTAGAAAACCTGTGCTTCATG 3273  
QY |||||  
DB 3554 ACGGCTTCTGTTAAAGTACAAAATTTAAATAATAGTAAAAATTCGCTCAATCACTACCAAGCCAG 3613  
DB |||||  
DB 3274 ACGGCTTCTGTTAAAGTACAAAATTTAAATAATAGTAAAAATTCGCTCAATCACTACCAAGCCAG 3333  
QY |||||  
DB 3614 GTAAAGCAAAAGGGGCTATTTTTCGCTATCCCTCAAAATCAAGCATGATTTGGCGGTCGTG 3673  
DB |||||  
DB 3334 GTAAAGCAAAAGGGGCTATTTTTCGCTATCCCTCAAAATCAAGCATGATTTGGCGGTCGTG 3393  
QY |||||  
DB 3674 GTGTGTTCTGACTTCCGAGGAGCGATTCAAGAAATCAAGATACATTTACACATTTGA 3733  
DB |||||  
DB 3394 GTGTGTTCTGACTTCCGAGGAGCGATTCAAGAAATCAAGATACATTTACACATTTGA 3453  
QY |||||  
DB 3734 CACCAACAGCTTTATCGTTATGGAACGTATGAGACGAAACCGTTTCATACAGAAAGGAC 3793  
DB |||||  
DB 3454 CACCAACAGCTTTATCGTTATGGAACGTATGAGACGAAACCGTTTCATACAGAAAGGAC 3513  
QY |||||  
DB 3794 ATTCGAAAAACATTTAAGACAAATCAATCTCTTTTATGATTTTGTATTTACACAGG 3853  
DB |||||  
DB 3514 ATTCGAAAAACATTTAAGACAAATCAATCTCTTTTATGATTTTGTATTTACACAGG 3573  
QY |||||  
DB 3854 CAAAAAGAACTATTTTACGAAAGCGATATTTTAAACAAACCGCTATTGATTTAGTTTATGC 3913  
DB |||||  
DB 3574 CAAAAAGAACTATTTTACGAAAGCGATATTTTAAACAAACCGCTATTGATTTAGTTTATGC 3633  
QY |||||  
DB 3914 CTATCTGATTTATCAAAATCTGATAAGGTTATCAAGCATTTTGTGTTTGAAGAGCCAG 3973  
DB |||||  
DB 3634 CTATCTGATTTATCAAAATCTGATAAGGTTATCAAGCATTTTGTGTTTGAAGAGCCAG 3693  
QY |||||  
DB 3974 TCTATGCTGACTTCAAAATCAGAAATTTAAATCTGCAAGAGCCAAATAATTTTCGAAA 4033  
DB |||||  
DB 3694 TCTATGCTGACTTCAAAATCAGAAATTTAAATCTGCAAGAGCCAAATAATTTTCGAAA 3753  
QY |||||  
DB 4034 ATATCCGAGAAATTTTTCGAAAGCTTTTGCAGTTGATCTAAGTGATCATTTTGGTA 4093  
DB |||||  
DB 3754 ATATCCGAGAAATTTTTCGAAAGCTTTTGCAGTTGATCTAAGTGATCATTTTGGTA 3813  
QY |||||  
DB 4094 TTGCTCGCATACCAAGAGCGGACATGAGAAATTTTGTGATCTTAATTTACCGTTTATCTT 4153  
DB |||||  
DB 3814 TTGCTCGCATACCAAGAGCGGACATGAGAAATTTTGTGATCTTAATTTACCGTTTATCTT 3873  
QY |||||  
DB 4154 TCAAGAAATGGAAGATTTGCTTTTCAAAACAAACAGATAAATAGGGCTTTTACTCGTTCAA 4213  
DB |||||  
DB 3874 TCAAGAAATGGAAGATTTGCTTTTCAAAACAAACAGATAAATAGGGCTTTTACTCGTTCAA 3933  
QY |||||  
DB 4214 GTCTAACCGTTTAAAGCGGTACAGAGGCAAAACAGATAGTCAACCTGTTTAAATC 4273  
DB |||||  
DB 3934 GTCTAACCGTTTAAAGCGGTACAGAGGCAAAACAGATAGTCAACCTGTTTAAATC 3993  
QY |||||  
DB 4274 TCTTATTTGACGAAACGAAATTTTTCAGGAGAAAAAGGGTTTAAATAGGGCGTAAATACGTTCA 4333

Db 3994 TCATTATGCGAAGCGAAATTTTCAGAGAAAAAGGGTTTAATAGGCGTAATAACGTC 4053  
Qy 4334 TGTATTACCTCTCTTTAGCCCTACTTTAGTTTCAAGGCTATTCAATCGAAACGTCGCAATATA 4393  
Db 4054 TGTATTACCTCTCTTTAGCCCTACTTTAGTTTCAAGGCTATTCAATCGAAACGTCGCAATATA 4113  
Qy 4394 ATATGTTTGAGTTTAAATATCGATTAGATCAACCCCTTAGAGAAAGAAAGATATCAAAA 4453  
Db 4114 ATATGTTTGAGTTTAAATATCGATTAGATCAACCCCTTAGAGAAAGAAAGATATCAAAA 4173  
Qy 4454 TTGTTAGAAGTGCCTATTTCAGAAAACCTATCAAGGGGCTAATAGGGAATACATTACCATTC 4513  
Db 4174 TTGTTAGAAGTGCCTATTTCAGAAAACCTATCAAGGGGCTAATAGGGAATACATTACCATTC 4233  
Qy 4514 TTTCGAAAGCTTGGGTATCAAGTGAATTAACCGATGAAGATTTATTTGTCGTCGCAAGGTT 4573  
Db 4234 TTTCGAAAGCTTGGGTATCAAGTGAATTAACCGATGAAGATTTATTTGTCGTCGCAAGGTT 4293  
Qy 4574 GGTTTAAATTAACAAAAAAGAGCGAAGCTCAACGTTGTCATTGTCAGATGGAAG 4633  
Db 4294 GGTTTAAATTAACAAAAAAGAGCGAAGCTCAACGTTGTCATTGTCAGATGGAAG 4353  
Qy 4634 AAGATTTAAGTGTATTATTTAGCGAAAAAAGCGATGTATCAAGCCCTTATTTAGTGACGA 4693  
Db 4354 AAGATTTAAGTGTATTATTTAGCGAAAAAAGCGATGTATCAAGCCCTTATTTAGTGACGA 4413  
Qy 4694 CCAAAAAGAGATTAGAGAGTGTCTAGGCAATTCCTGACGACATTTAGATTAATTCCTGA 4753  
Db 4414 CCAAAAAGAGATTAGAGAGTGTCTAGGCAATTCCTGACGACATTTAGATTAATTCCTGA 4473  
Qy 4754 AGGTACTGAAGGCGAATCAGGAAATTTTCTTTAAGATTAAACCCAGGAAGAAATGCTGGCA 4813  
Db 4474 AGGTACTGAAGGCGAATCAGGAAATTTTCTTTAAGATTAAACCCAGGAAGAAATGCTGGCA 4533  
Qy 4814 TTCAACTTGTAGTGTAAATCATTTGTGCTATCGATCATTAAAGTAAATAAAGAAAGAA 4873  
Db 4534 TTCAACTTGTAGTGTAAATCATTTGTGCTATCGATCATTAAAGTAAATAAAGAAAGAA 4593  
Qy 4874 AAGAAAGCTATATAAAGGCGCTCAGCAATTTCTTTTGAATTTAGCTTTAGAGCATACATTCATTCAG 4933  
Db 4594 AAGAAAGCTATATAAAGGCGCTCAGCAATTTCTTTTGAATTTAGCTTTAGAGCATACATTCATTCAG 4653  
Qy 4934 AGACTTTAAACAAGCTAGCAGAACGCCCTTAAACCGGACACACAACTCGATTTGTTAGCT 4993  
Db 4654 AGACTTTAAACAAGCTAGCAGAACGCCCTTAAACCGGACACACAACTCGATTTGTTAGCT 4713  
Qy 4994 ATGATACAGCTGAAATTAACCCGCACTATGCCATTACATTTATATCTATGATACGTCG 5053  
Db 4714 ATGATACAGCTGAAATTAACCCGCACTATGCCATTACATTTATATCTATGATACGTCG 4773  
Qy 5054 TTTGTTTTTTCTTTGCTGTTTAGCGAATGATTAGCAGAAATATACAGAGTAAGATTTTAA 5113  
Db 4774 TTTGTTTTTTCTTTGCTGTTTAGCGAATGATTAGCAGAAATATACAGAGTAAGATTTTAA 4833  
Qy 5114 TTAATTTATTAGGGGAGAGAGAGTAGTACCGGAAACCTTTTAGTTGGCTTGGACTGAA 5173  
Db 4834 TTAATTTATTAGGGGAGAGAGAGTAGTACCGGAAACCTTTTAGTTGGCTTGGACTGAA 4893  
Qy 5174 CGAAGTGAGGGAAGGCTACTAAACCTCGAGGGCAGTGAGAGCGAAGCGAAGCACTTGA 5233  
Db 4894 CGAAGTGAGGGAAGGCTACTAAACCTCGAGGGGCGAGTGAGAGCGAAGCGAAGCACTTGA 4953  
Qy 5234 TTTTAAATTTCTTATCTTTATAGTCAATTAGAGPATATCTTATTGTCCTATAAATAT 5293  
Db 4954 TTTTAAATTTCTTATCTTTATAGTCAATTAGAGPATATCTTATTGTCCTATAAATAT 5013  
Qy 5294 TTAGCAGCATATAGATTTTATGAATGCTATTTAGTTAGTTCAGTATTTAGAGGAGGAA 5353  
Db 5014 TTAGCAGCATATAGATTTTATGAATGCTATTTAGTTAGTTCAGTATTTAGAGGAGGAA 5073  
Qy 5354 ATCTTCGAGAAATTTTGAAGAACCCGATTACATGGAATGGAATAGTTAGTTCTTCTGCTTACG 5413  
Db 5074 ATCTTCGAGAAATTTTGAAGAACCCGATTACATGGAATGGAATAGTTAGTTCTTCTGCTTACG 5133

Qy 5414 TGGTTTTTAACATAAAGTAGTAGTAATTTTGAATTTTGGTGTGCTGCTCTTGTGTTAGTA 5473  
Db 5134 TGGTTTTTAACATAAAGTAGTAGTAATTTTGAATTTTGGTGTGCTGCTCTTGTGTTAGTA 5193  
Qy 5474 TTTGCTAGTCAAAAGTAGTAATAAATA 5497  
Db 5194 TTTGCTAGTCAAAAGTAGTAATAAATA 5217  
RESULT 4  
AR381899 LOCUS AR381899 5230 bp DNA linear PAT 18-DEC-2003  
DEFINITION Sequence 26 from patent US 6610300.  
ACCESSION AR381899  
VERSION AR381899.1 GI:40090248  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 5230)  
AUTHORS Segers,R.P.A.M., Waterfield,N.R., Frandsen,P.L. and Wells,J.M.  
TITLE Clostridium perfringens vaccine  
JOURNAL Patent: US 6610300-A 26 26-AUG-2003;  
FEATURES  
source  
1. 5230  
/organism="unknown"  
/mol\_type="genomic DNA"  
ORIGIN  
Query Match 90.8%; Score 4993.4; DB 6; Length 5230;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 4997; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
Qy 495 TCTAACTAGTAGTCCGCTGCTAACAAAGCCGAAAGAGAGAGTGGTGGCTGCTGCCA 554  
Db 238 TCTCTAGAATGATCCGCTGCTAACAAAGCCGAAAGAGAGTGGTGGCTGCTGCCA 287  
Qy 555 CCGCTGACCAATTAAGTACATACATCCCTTGGGGCTCTAAACGGGTCTTGAAGGGTTTTT 614  
Db 288 CCGCTGACCAATTAAGTACATACATCCCTTGGGGCTCTAAACGGGTCTTGAAGGGTTTTT 347  
Qy 615 TGCTGAAAGGAGGAATATATCCGGATGACCTCGAGGCAAGCTCTAGAATCGATACGATT 674  
Db 348 TGCTGAAAGGAGGAATATATCCGGATGACCTCGAGGCAAGCTCTAGAATCGATACGATT 407  
Qy 675 TTGAAGTGGCAACAGATATAAAGAGAGTGTAAATTTGTTGCTGAACTTTTAAACAAG 734  
Db 408 TTGAAGTGGCAACAGATATAAAGAGAGTGTAAATTTGTTGCTGAACTTTTAAACAAG 467  
Qy 735 CAATATACATCATTTGTCGACAGATAGCGACAGAGAGGCGAAACATTCGCTGGTCGA 794  
Db 468 CAATATACATCATTTGTCGACAGATAGCGACAGAGAGGCGAAACATTCGCTGGTCGA 527  
Qy 795 TCATTCTATAAGCAAAATGCCCTTTCTAAAGATAAAGAGAGTATAAAGAGCTATGATCAATA 854  
Db 528 TCATTCTATAAGCAAAATGCCCTTTCTAAAGATAAAGAGAGTATAAAGAGCTATGATCAATA 587  
Qy 855 GTTTAGAAAAGATGTATCGTAGCGGTTTCAAAATTTGCAACCGAGGAATGAATTAAT 914  
Db 588 GTTTAGAAAAGATGTATCGTAGCGGTTTCAAAATTTGCAACCGAGGAATGAATTAAT 647  
Qy 915 ATCCCTTTTATCAAGAGCGCAAAAGAAAAACGAAATGATACCAATCAGTGCAGAAAAA 974  
Db 648 ATCCCTTTTATCAAGAGCGCAAAAGAAAAACGAAATGATACCAATCAGTGCAGAAAAA 707  
Qy 975 AGATATAATGGGAGATGAAGCGGTTGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1034  
Db 708 AGATATAATGGGAGATGAAGCGGTTGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 767  
Qy 1035 CGAAACAGCAAAAGATGCGCGAAACGTAAGAGAGTTATCGAATAGAGCTTAGAGAGCA 1094  
Db 768 CGAAACAGCAAAAGATGCGCGAAACGTAAGAGAGTTATCGAATAGAGCTTAGAGAGCA 827



QY	1095	ACTTAAGAGTGTGTGATGTCAGTATCTTAAAAATTTTGTATATAATAGGAATTTGAAGTTA	1154	QY	2175	GGTTTCAAAAAGGTTTAAATAAGAGAGCAATCAATGCAATAGCTAGCTAGTATATATTTT	2234
DB	828	ACTTAAGAGTGTGTGATGTCAGTATCTTAAAAATTTTGTATATAATAGGAATTTGAAGTTA	887	DB	1908	GGTTTCAAAAAGGTTTAAATAAGAGAGCAATCAATGCAATAGCTAGCTAGTATATATTTT	1967
QY	1155	AATTAGATGCTAAAAATTTGTATTAAGAGAGGATGATACATGAACAAAAATATAAAAT	1214	QY	2235	TTGGACAACTGGAGAAATTTAGAGAACTGCTCTCCAAAGACCAGTTACAAAGAGCTAGTG	2294
DB	888	AATTAGATGCTAAAAATTTGTATTAAGAGAGGATGATACATGAACAAAAATATAAAAT	947	DB	1968	TTGGACAACTGGAGAAATTTAGAGAACTGCTCTCCAAAGACCAGTTACAAAGAGCTAGTG	2027
QY	1215	ATTCTCAAAAATTTTAAACGAGTGAAAAAGTACTCAACCAAAATAAATAAAACAATGAAAT	1274	QY	2295	CACATAACATATATTAACGCTATTAAGTGTGTGGAACACCTGTATATATATGAAAAACCG	2354
DB	948	ATTCTCAAAAATTTTAAACGAGTGAAAAAGTACTCAACCAAAATAAATAAAACAATGAAAT	1007	DB	2028	CACATAACATATATTAACGCTATTAAGTGTGTGGAACACCTGTATATATATGAAAAACCG	2087
QY	1275	TAAAGAAACCGATACCGTTTACGAATTTGGAACAGGTAAAGGCATTTAAACGACGAAC	1334	QY	2355	TAGAAGAAATTAAGAGCAAGAGAGAAATTTAGAGAAATTTAATGCCATATGCGTGGCGGT	2414
DB	1008	TAAAGAAACCGATACCGTTTACGAATTTGGAACAGGTAAAGGCATTTAAACGACGAAC	1067	DB	2088	TAGAAGAAATTAAGAGCAAGAGAGAAATTTAGAGAAATTTAATGCCATATGCGTGGCGGT	2147
QY	1335	TGGCTAAAAATAAGTAAACAGGTAAACGCTATTTGAATTTAGACAGTCACTTATCAACTTAT	1394	QY	2415	TAGGATGGGAACATATCAATTTCTCGAGAAATACAAATTTGAAGGATTAACATGACACTG	2474
DB	1068	TGGCTAAAAATAAGTAAACAGGTAAACGCTATTTGAATTTAGACAGTCACTTATCAACTTAT	1127	DB	2148	TAGGATGGGAACATATCAATTTCTCGAGAAATACAAATTTGAAGGATTAACATGACACTG	2207
QY	1395	CGTCAGAAAAATTAACACTGAATCTCGTGTCACTTTAATTCACCAAGATATTTCTACAGT	1454	QY	2475	GGCAATGAATTTAGCTCTTTTACGTATATAAGAGCCGTTTTTATCTTAATATACCGCT	2534
DB	1128	CGTCAGAAAAATTAACACTGAATCTCGTGTCACTTTAATTCACCAAGATATTTCTACAGT	1187	DB	2208	GGCAATGAATTTAGCTCTTTTACGTATATAAGAGCCGTTTTTATCTTAATATACCGCT	2267
QY	1455	TTCAATTCCTTAACAAACAGAGGTATATAATTTGGAGTATTCCTTACCATTTAAGCA	1514	QY	2535	CTTTTATAGAAAAATCTTAGCGTGGTTTTTTTCCGAAATGCTGGCGGTACCCCAAGA	2594
DB	1188	TTCAATTCCTTAACAAACAGAGGTATATAATTTGGAGTATTCCTTACCATTTAAGCA	1247	DB	2268	CTTTTATAGAAAAATCTTAGCGTGGTTTTTTTCCGAAATGCTGGCGGTACCCCAAGA	2327
QY	1515	CACAAATTAATTAAGAGTGTGTTTTGAAAGCCATGGCTGACATCTATCTGATTTGTTG	1574	QY	2595	ATTGAAATGAGTAGATCAAAATTTATTCAGGAATAGAAATCAGGAAAAATCAGATCAACCAT	2654
DB	1248	CACAAATTAATTAAGAGTGTGTTTTGAAAGCCATGGCTGACATCTATCTGATTTGTTG	1307	DB	2328	ATTGAAATGAGTAGATCAAAATTTATTCAGGAATAGAAATCAGGAAAAATCAGATCAACCAT	2387
QY	1575	AAGAGGATTTCAAGGTAACCTTGATATTTCCAGAACACTAGGTTGCTTTGACACA	1634	QY	2655	AAAAACACTAGAACCAATTTGCAAAAGTTAACTCAACGCTAGTAGTGGATTTAAATCCC	2714
DB	1308	AAGAGGATTTCAAGGTAACCTTGATATTTCCAGAACACTAGGTTGCTTTGACACA	1367	DB	2388	AAAAACACTAGAACCAATTTGCAAAAGTTAACTCAACGCTAGTAGTGGATTTAAATCCC	2447
QY	1635	CTCAAGTCTCGAATTCAGCAATTTGCTTAAGCTGCCAGCGGAATGCTTTCACTTAAACCA	1694	QY	2715	AAATCAGCCCAACAGAACAGAGCCAGAAACAGAAATCAGAACCAAGTAACATTTGGATTTAGA	2774
DB	1368	CTCAAGTCTCGAATTCAGCAATTTGCTTAAGCTGCCAGCGGAATGCTTTCACTTAAACCA	1427	DB	2448	AAATCAGCCCAACAGAACAGAGCCAGAAACAGAAATCAGAACCAAGTAACATTTGGATTTAGA	2507
QY	1695	AAGTAACAGTGTCTTAATAAACCCTTACCCGCAATACACAGATGTTCCAGATAAATATT	1754	QY	2775	AATGGAAGAGAAAAAAGCAATGACTTCGTGTGAATTAATGACGAAATCGTGTCTATTTT	2834
DB	1428	AAGTAACAGTGTCTTAATAAACCCTTACCCGCAATACACAGATGTTCCAGATAAATATT	1487	DB	2508	AATGGAAGAGAAAAAAGCAATGACTTCGTGTGAATTAATGACGAAATCGTGTCTATTTT	2567
QY	1755	GGAGCTATATACGTACTTTGTTTCAAAATGGGTCAATCGAGAAATATCGTCAACTGTTTA	1814	QY	2835	TTTTTAAAGCGGTATCTAGATATAAGAAACAAACGAACTGTAATAAGAAACGAAAAAAG	2894
DB	1488	GGAGCTATATACGTACTTTGTTTCAAAATGGGTCAATCGAGAAATATCGTCAACTGTTTA	1547	DB	2568	TTTTTAAAGCGGTATCTAGATATAAGAAACAAACGAACTGTAATAAGAAACGAAAAAAG	2627
QY	1815	CTAAAAATCAGTTTCATCAAGCAATGAACACGCGCAAGATGAACAAATTAAGTACCGTTA	1874	QY	2895	AGCCATGACACATTTTATAAAATGTTTGAACGACATTTTATAATGATAGCCGATAAGAT	2954
DB	1548	CTAAAAATCAGTTTCATCAAGCAATGAACACGCGCAAGATGAACAAATTAAGTACCGTTA	1607	DB	2628	AGCCATGACACATTTTATAAAATGTTTGAACGACATTTTATAATGATAGCCGATAAGAT	2687
QY	1875	CTTATGAGCAATGTTGCTATTTTATAGTTATCTATTTTAAACGGAGGAAATAAT	1934	QY	2955	TGCGAAACCAACGCTTATCAGTTAGTACAGTAACTCTTCCCTCGTAAGAGTATTTTAA	3014
DB	1608	CTTATGAGCAATGTTGCTATTTTATAGTTATCTATTTTAAACGGAGGAAATAAT	1667	DB	2688	TGCGAAACCAACGCTTATCAGTTAGTACAGTAACTCTTCCCTCGTAAGAGTATTTTAA	2747
QY	1935	TCTATGAGTCGCTTTGTAATTTGGAAGATTAACGTTACTAAAGGGAATGTAGATAAA	1994	QY	3015	TTAACTTTGTTTGAAGACGCTATATAACCGTACTATCATTTATATAGGGAATCAGAGAT	3074
DB	1668	TCTATGAGTCGCTTTGTAATTTGGAAGATTAACGTTACTAAAGGGAATGTAGATAAA	1727	DB	2748	TTAACTTTGTTTGAAGACGCTATATAACCGTACTATCATTTATATAGGGAATCAGAGAT	2807
QY	1995	TTATTAGGTATCTACTGACGCTTCCAGAGGCTTAAAGAGTCCCTAGCGCTCTTATCA	2054	QY	3075	TTTCAAGATCTAAGCTACTGAATTTAAGAAATTTGTTAGCAATCAATCGGAATCGTTTG	3134
DB	1728	TTATTAGGTATCTACTGACGCTTCCAGAGGCTTAAAGAGTCCCTAGCGCTCTTATCA	1787	DB	2808	TTTCAAGATCTAAGCTACTGAATTTAAGAAATTTGTTAGCAATCAATCGGAATCGTTTG	2867
QY	2055	TGGGAAGCTCGGATCATATGACACAPAAATAAATCTCGCAACAGCACTTCGGAAGATGG	2114	QY	3135	ATTGCTTTTTTTGTTATTTTATAGAGGTCGAGTTTGTATGAATCATGTATGAATGTAA	3194
DB	1788	TGGGAAGCTCGGATCATATGACACAPAAATAAATCTCGCAACAGCACTTCGGAAGATGG	1847	DB	2868	ATTGCTTTTTTTGTTATTTTATAGAGGTCGAGTTTGTATGAATCATGTATGAATGTAA	2927
QY	2115	GACGAATCGAANAACCTTTTACGCTGGATTACATATCTAATAAAGCCGTTAAGGAGAC	2174	QY	3195	AACCTTATAAAAAATAGTTTTTATGGAGATAAGAAATTTAGCAAAATCTATACACTAGA	3254
DB	1848	GACGAATCGAANAACCTTTTACGCTGGATTACATATCTAATAAAGCCGTTAAGGAGAC	1907	DB	2928	AACCTTATAAAAAATAGTTTTTATGGAGATAAGAAATTTAGCAAAATCTATACACTAGA	2987







Db	1908	GGGTTCAAAGAGGTTTAAATTAAGGAGAGCAACTCAATGCAATTAGCTAGCACTATATTTT	1967
QY	2235	TTGGACAACGTTGGAGAAATTTAGAGAACGTTGCTCTCCAAAGACGAGTTTACAAAGAGCTAGTG	2294
Db	1968	TTGGACAACGTTGGAGAAATTTAGAGAACGTTGCTCTCCAAAGACGAGTTTACAAAGAGCTAGTG	2027
QY	2295	CACATAACATAATTAATTAACCGCTATAAGCTGTGTGGAAACACCTGTTATATATGAAAAAGCCG	2354
Db	2028	CACATAACATAATTAATTAACCGCTATAAGCTGTGTGGAAACACCTGTTATATATGAAAAAGCCG	2087
QY	2355	TAGAAAGAAATTAAGAACCAAGAGAGAGAAATTTAGAGAGAGATTTTAATGCCATATGCGTGCCGCT	2414
Db	2088	TAGAAAGAAATTAAGAACCAAGAGAGAGAAATTTAGAGAGAGATTTTAATGCCATATGCGTGCCGCT	2147
QY	2415	TAGGATGGGAACATATCAATCTTTCTTGGAGAAATCAAAATTTGAAGGATTCATGACACATG	2474
Db	2148	TAGGATGGGAACATATCAATCTTTCTTGGAGAAATCAAAATTTGAAGGATTCATGACACATG	2507
QY	2475	GGCAAAATGAAATTTAGTCTCTTTTACGTATAAAGAGCCGTTTATTTCTTAATATAACGGCT	2534
Db	2208	GGCAAAATGAAATTTAGTCTCTTTTACGTATAAAGAGCCGTTTATTTCTTAATATAACGGCT	2267
QY	2535	CTTTTTTATAGAAAAAATCTTTAGCGTGGTTTTTTTCGGAATGCTGCGGTACCCCAAGA	2594
Db	2268	CTTTTTTATAGAAAAAATCTTTAGCGTGGTTTTTTTCGGAATGCTGCGGTACCCCAAGA	2327
QY	2595	ATTAGAAATGATGATGATCAAAATTAATCACCAGATAGAAATCAGGAAATCAGATCCACCAT	2654
Db	2328	ATTAGAAATGATGATGATCAAAATTAATCACCAGATAGAAATCAGGAAATCAGATCCACCAT	2387
QY	2655	AAAAACACTAGAAACAAATTCGAAAGTTAACTTAACCTAACCGCTAGTGTAGTATTAATCCG	2714
Db	2388	AAAAACACTAGAAACAAATTCGAAAGTTAACTTAACCTAACCGCTAGTGTAGTATTAATCCG	2447
QY	2715	AAATGAGCCACAGAAACAGAGCCAGAAAACAGAAATCAGAACAGTAACTTGAAATTTAGA	2774
Db	2448	AAATGAGCCACAGAAACAGAGCCAGAAAACAGAAATCAGAACAGTAACTTGAAATTTAGA	2507
QY	2775	AATGGAAGAAAGAAAAAGCAATGACTTCGTGTGAATTAATGCACGAAATCGTGCTTATTTT	2834
Db	2508	AATGGAAGAAAGAAAAAGCAATGACTTCGTGTGAATTAATGCACGAAATCGTGCTTATTTT	2567
QY	2835	TTTTTTTTAAAGCCGTATACCTAGATATTAACGAAACAGAACTGAATAGAAAAAG	2894
Db	2568	TTTTTTTTAAAGCCGTATACCTAGATATTAACGAAACAGAACTGAATAGAAAAAG	2627
QY	2895	AGCCATGACACATTTATAAAATGTTTTGACGACATTTTATAAATGCAATGCCCGGATAAGAT	2954
Db	2628	AGCCATGACACATTTATAAAATGTTTTGACGACATTTTATAAATGCAATGCCCGGATAAGAT	2687
QY	2955	TGCGAAACCAACCGTTATCAAGTTAGTCAGATGAACCTCTTCCCTCGTGAAGTTATTTAA	3014
Db	2688	TGCGAAACCAACCGTTATCAAGTTAGTCAGATGAACCTCTTCCCTCGTGAAGTTATTTAA	2747
QY	3015	TTAACTTTGTTTGAAGACGCTATATACCGCTACTATCATATATAGGGAATCAGAGAT	3074
Db	2748	TTAACTTTGTTTGAAGACGCTATATACCGCTACTATCATATATAGGGAATCAGAGAT	2807
QY	3075	TTTCAAGTATCTAAGCTACTGAAATTTAAGAAATTTGTTAAGCAATCAATCGGAATCGTTTG	3134
Db	2808	TTTCAAGTATCTAAGCTACTGAAATTTAAGAAATTTGTTAAGCAATCAATCGGAATCGTTTG	2867
QY	3135	ATTGCTTTTTTTGTAATCATTTATAGAAAGTGAGTTTCTATGAATCATGATGAATGTAA	3194
Db	2868	ATTGCTTTTTTTGTAATCATTTATAGAAAGTGAGTTTCTATGAATCATGATGAATGTAA	2927
QY	3195	AACCTATATAAAAAATAGTTTATTTGAGATTAAGAAAAATTAGCAAAATCTATPACACTAGA	3254
Db	2928	AACCTATATAAAAAATAGTTTATTTGAGATTAAGAAAAATTTAGCAAAATCTATPACACTAGA	2987
QY	3255	AAAGCTTTAAGAAAGAGTTAGAAAAAGAAATATCTACTTTAGAAAAAATAATCAGATAAGTA	3314
Db	2988	AAAGCTTTAAGAAAGAGTTAGAAAAAGAAATATCTACTTTAGAAAAAATAATCAGATAAGTA	3047



|||||  
Db 895 ATGCTRAAAATTTGTAATTAGAGAGGAGTATACATGACCAAAAATATAAAATATTCTC 954  
QY 1221 AAAAATTTTAAACGAGTGAAAGAGTCTCAACCAAAATATAAAATTTAAAG 1280  
Db 955 AAAAATTTTAAACGAGTGAAAGAGTCTCAACCAAAATATAAAATTTAAAG 1014  
QY 1281 AAAACCGATACCGTTTACGAAATTTGGAACAGGTAAAGGCGATTTAAACGAAATCGGCTA 1340  
Db 1015 AAAACCGATACCGTTTACGAAATTTGGAACAGGTAAAGGCGATTTAAACGAAATCGGCTA 1074  
QY 1341 AAAAATTAACAGAGTAAACGCTCTATTGAATTTAGACAGTCTATCTCAATCTATCGTCAG 1400  
Db 1075 AAAAATTAACAGAGTAAACGCTCTATTGAATTTAGACAGTCTATCTCAATCTATCGTCAG 1134  
QY 1401 AAAAATTAACAGTAAATCTCGTCTCAATTTAAATTCACCAAGATATTTCTACAGTTCAAT 1460  
Db 1135 AAAAATTAACAGTAAATCTCGTCTCAATTTAAATTCACCAAGATATTTCTACAGTTCAAT 1194  
QY 1461 TCCCTAAACAAACAGAGGTATAAAATTTGCGGAGTATTCCTTACCATTAAAGCACACAA 1520  
Db 1195 TCCCTAAACAAACAGAGGTATAAAATTTGCGGAGTATTCCTTACCATTAAAGCACACAA 1254  
QY 1521 TTATTAAAAAGTGGTTTTTGAAGCCATCGGCTCGACATCTATCTGATTTGAAGAAG 1580  
Db 1255 TTATTAAAAAGTGGTTTTTGAAGCCATCGGCTCGACATCTATCTGATTTGAAGAAG 1314  
QY 1581 GATTCTACAGCGTACCTTCGATTAATTCACCAAGTCTAGGTTGCTTGCACACTCAAG 1640  
Db 1315 GATTCTACAGCGTACCTTCGATTAATTCACCAAGTCTAGGTTGCTTGCACACTCAAG 1374  
QY 1641 TCTCGATTCAGCAATTTGCTTAACTGCCAGCGGAATGCTTTCATCTCAACCAAAAGTAA 1700  
Db 1375 TCTCGATTCAGCAATTTGCTTAACTGCCAGCGGAATGCTTTCATCTCAACCAAAAGTAA 1434  
QY 1701 ACAGTCTCTTAATAAAGTAACTCCGCCATACCAAGATGTTCCAGATAAATTTGAAGC 1760  
Db 1435 ACAGTCTCTTAATAAAGTAACTCCGCCATACCAAGATGTTCCAGATAAATTTGAAGC 1494  
QY 1761 TATATACGTAATTTGTTTCAAAATGGGTCAATTCGAGATATCGTCAACTGTTTACTAAA 1820  
Db 1495 TATATACGTAATTTGTTTCAAAATGGGTCAATTCGAGATATCGTCAACTGTTTACTAAA 1554  
QY 1821 ATCAGTTTTCATCAGCAATTAACAGCGCAAGTAAACAAATTAAGTACCGTACTTATG 1880  
Db 1555 ATCAGTTTTCATCAGCAATTAACAGCGCAAGTAAACAAATTAAGTACCGTACTTATG 1614  
QY 1881 AGCAAGTATTTGCTATTTTAAATAGTTATCTATTTAAACGAGGGAATTAATTTCTATG 1940  
Db 1615 AGCAAGTATTTGCTATTTTAAATAGTTATCTATTTAAACGAGGGAATTAATTTCTATG 1674  
QY 1941 AGTCGCTTTTGTAAATTTGGAAGTTTACAGTTTAAAGGGAATGTAGATAAATTTA 2000  
Db 1675 AGTCGCTTTTGTAAATTTGGAAGTTTACAGTTTAAAGGGAATGTAGATAAATTTA 1734  
QY 2001 GGTATACTACTGACAGCTTCCAGAGCTTAAGAGGTCCTAGCGCTCTTATCATGGGA 2060  
Db 1735 GGTATACTACTGACAGCTTCCAGAGCTTAAGAGGTCCTAGCGCTCTTATCATGGGA 1794  
QY 2061 AGCTCGGATCATATGCAAGCAAAATTAACCTCGCAACAGCACTTCGAGAAATGGAGCAA 2120  
Db 1795 AGCTCGGATCATATGCAAGCAAAATTAACCTCGCAACAGCACTTCGAGAAATGGAGCAA 1854  
QY 2121 TCGAGAAACCTCTTTAGCTGGATTACATCTTAATTAAGCCGTAAAGAGCGGTTTC 2180  
Db 1855 TCGAGAAACCTCTTTAGCTGGATTACATCTTAATTAAGCCGTAAAGAGCGGTTTC 1914  
QY 2181 AAAAGGTTTAAATAAGGAGAGCAATCAATGCAATTTAGCTAGACTATTTTTTGGAC 2240  
Db 1915 AAAAGGTTTAAATAAGGAGAGCAATCAATGCAATTTAGCTAGACTATTTTTTGGAC 1974  
QY 2241 AACGTGGAGATTTAGAGAACTGCTCTCCAGACCAAGTAAACAGAGCTAGTGCCTAA 2300  
|||||

Db 1975 AACGTGGAGATTTAGAGAACTGCTCTCCAAAGACCAGTTACAAAGAGCTAGTGCACTAA 2034  
QY 2301 ACATATTTTAAACGCTATAAGCTGTGGAACACATGTATATATGGAAGCCGTAGAG 2360  
Db 2035 ACATATTTTAAACGCTATAAGCTGTGGAACACATGTATATGGAAGCCGTAGAG 2094  
QY 2361 AATTAAAGCAAGAGGAGAAATTTAGAGAGATTTAATGCCATATGCGTGGCGTTAGGAT 2420  
Db 2095 AATTAAAGCAAGAGGAGAAATTTAGAGAGATTTAATGCCATATGCGTGGCGTTAGGAT 2154  
QY 2421 GGGAAACATATCAATTTCTTGGAGAAATCAAAATTTGAAGGATTTACATGACATGGGCA 2480  
Db 2155 GGGAAACATATCAATTTCTTGGAGAAATCAAAATTTGAAGGATTTACATGACATGGGCA 2214  
QY 2481 TGAATTTACGCTCTTTACGCTATAAAGAGCCGTTTATTTCTTAATATACCGCTCTTTT 2540  
Db 2215 TGAATTTACGCTCTTTACGCTATAAAGAGCCGTTTATTTCTTAATATACCGCTCTTTT 2274  
QY 2541 ATAGAAAAATTCCTTTAGCGTGGTTTTTTTTCGAAATGCTGGCGGTACCCCAAGAAATAGA 2600  
Db 2275 ATAGAAAAATTCCTTTAGCGTGGTTTTTTTTCGAAATGCTGGCGGTACCCCAAGAAATAGA 2334  
QY 2601 AATGAGTAGATCAAAATTTTCAAGATTAAGTAACTAGATCAGGAAAAATCAGATCCAAATAA 2660  
Db 2335 AATGAGTAGATCAAAATTTTCAAGATTAAGTAACTAGATCAGGAAAAATCAGATCCAAATAA 2394  
QY 2661 ACTAGAACAAAATGCAAAAGTTAACTAATCAACCTAGTGTAGTATTTAATCCCAATAGA 2720  
Db 2395 ACTAGAACAAAATGCAAAAGTTAACTAATCAACCTAGTGTAGTATTTAATCCCAATAGA 2454  
QY 2721 GCCAACAGAACAGAGCCAGAAACAGAAATCAGAAATCAGAAATCAGAAATCAGAAATCAGAA 2780  
Db 2455 GCCAACAGAACAGAGCCAGAAACAGAAATCAGAAATCAGAAATCAGAAATCAGAAATCAGAA 2514  
QY 2781 AGAAGAAAAAGCAATGACTTCTGTTGTAATTAATCAGCAAAATCGTTGCTTATTTTTTTT 2840  
Db 2515 AGAAGAAAAAGCAATGACTTCTGTTGTAATTAATCAGCAAAATCGTTGCTTATTTTTTTT 2574  
QY 2841 AAAAGCGGTATCTAGATATAACCAACAGAACTGAATAGAAACAGAAAGAGCCAT 2900  
Db 2575 AAAAGCGGTATCTAGATATAACCAACAGAACTGAATAGAAACAGAAAGAGCCAT 2634  
QY 2901 GACACATTTATAAATGTTTACGACATTTTATAAATGCAATAGCCCATAGATTTGCCAA 2960  
Db 2635 GACACATTTATAAATGTTTACGACATTTTATAAATGCAATAGCCCATAGATTTGCCAA 2694  
QY 2961 ACCAACCGTTATCAGTTAGTCAGATGAATCTTCCCTCGTAAGAGTATTTAATTAAC 3020  
Db 2695 ACCAACCGTTATCAGTTAGTCAGATGAATCTTCCCTCGTAAGAGTATTTAATTAAC 2754  
QY 3021 TTGTTTCAAGACCGGTATATAACCGTACTATCATTTATAGGAAATCGTTTGAATGCT 3140  
Db 2755 TTGTTTCAAGACCGGTATATAACCGTACTATCATTTATAGGAAATCGTTTGAATGCT 2814  
QY 3081 GTATCTAAGCTACTGAATTTAAGAAATTTGAACAAATCAATCGGAAATCGTTTGAATGCT 3140  
Db 2815 GTATCTAAGCTACTGAATTTAAGAAATTTGAACAAATCAATCGGAAATCGTTTGAATGCT 2874  
QY 3141 TTTTGTATTTCAATTTATAGAGGTGGAGTTTGTATGAATCATGATGAATGTAAAACTTA 3200  
Db 2875 TTTTGTATTTCAATTTATAGAGGTGGAGTTTGTATGAATCATGATGAATGTAAAACTTA 2934  
QY 3201 TATAAAAAATAGTTTATTTGGAGATAAGAAAAATAGCAAAATATCTATACACTAGAAAGCTT 3260  
Db 2935 TATAAAAAATAGTTTATTTGGAGATAAGAAAAATAGCAAAATATCTATACACTAGAAAGCTT 2994  
QY 3261 TAGAAAGAGTTAGAAAAAGAAATATCTATTAGAAAAAATAAGAAATCAGATAAGTATTTTC 3320  
Db 2995 TAGAAAGAGTTAGAAAAAGAAATATCTATTAGAAAAAATAAGAAATCAGATAAGTATTTTC 3054  
QY 3321 TTCCGAGGGGGAAGATTTATATATAGTTTAAATAGAAAAATAAGAAATATTTATTCGAT 3380  
Db 3055 TTCCGAGGGGGAAGATTTATATATAGTTTAAATAGAAAAATAAGAAATATTTATTCGAT 3114







[illegible]



Db 4195 AAGTGCTATTACAGAAAATCTATCAAGGGCTAATAGGGAATACATTTACCATCTTTTGCAA 4254  
QY 4521 AGCTTGGGTATCAAGTGATTTAAACAGTAAGATTTATTTGTCGGTCAAGGGTGGTTTAA 4580  
Db 4255 AGCTTGGGTATCAAGTGATTTAAACAGTAAGATTTATTTGTCGGTCAAGGGTGGTTTAA 4314  
QY 4581 ATTCAAGAAAAAAGAGGCAAGCTCAACGTGTTTCATTTTGTGAGATGGAAGAAGATT 4640  
Db 4315 ATTCAAGAAAAAAGAGGCAAGCTCAACGTGTTTCATTTTGTGAGATGGAAGAAGATT 4374  
QY 4641 AATGGCTTATATTAGCGAAAAAGCGATGATATCAAGCCTTATTTAGTGACCAACAAAA 4700  
Db 4375 AATGGCTTATATTAGCGAAAAAGCGATGATATCAAGCCTTATTTAGTGACCAACAAAA 4434  
QY 4701 AGAGATTAGAGAGTCTAGGCAATTCCTGAAACGACATTAGATATATTTCTGAGGTACT 4760  
Db 4435 AGAGATTAGAGAGTCTAGGCAATTCCTGAAACGACATTAGATATATTTCTGAGGTACT 4494  
QY 4761 GAAGGCGAATCAGAAATTTCTTTTAAGATTTAAACAGGAAGAAATGGTGGCATTCACT 4820  
Db 4495 GAAGGCGAATCAGAAATTTCTTTTAAGATTTAAACAGGAAGAAATGGTGGCATTCACT 4554  
QY 4821 TGCTAGTGTAAATCATTTGCTATCGATCATTAAGTAAAAAAGAAAGAAAGAAAG 4880  
Db 4555 TGCTAGTGTAAATCATTTGCTATCGATCATTAAGTAAAAAAGAAAGAAAGAAAG 4614  
QY 4881 CTATATAAGCGCTGACAAATTTCTTTTGAATTTAGAGCATACATTCATTCAAGAGACTTT 4940  
Db 4615 CTATATAAGCGCTGACAAATTTCTTTTGAATTTAGAGCATACATTCATTCAAGAGACTTT 4674  
QY 4941 AAACAAGCTAGAGAGCGCCTAAACGGACACACAACTCGATTTGTTTAGCTATGATAC 5000  
Db 4675 AAACAAGCTAGAGAGCGCCTAAACGGACACACAACTCGATTTGTTTAGCTATGATAC 4734  
QY 5001 AGCTGAAATTAACACCGCACTATGCCATTATCATTTATATCATGATACGTTGTTT 5060  
Db 4735 AGCTGAAATTAACACCGCACTATGCCATTATCATTTATATCATGATACGTTGTTT 4794  
QY 5061 TTTCTTTGCTTTTAGCGAATGATTAGCAGAAATATACAGAGTAAGATTTTAAATTA 5120  
Db 4795 TTTCTTTGCTTTTAGCGAATGATTAGCAGAAATATACAGAGTAAGATTTTAAATTA 4854  
QY 5121 TTAGGGGAGAGAGAGTAGCCGAAACCTTTTAGTTGGCTTGGCTGACGACGAAGTG 5180  
Db 4855 TTAGGGGAGAGAGAGTAGCCGAAACCTTTTAGTTGGCTTGGCTGACGACGAAGTG 4914  
QY 5181 AGGCAAGGCTACTAAAACTGAGGGGAGTGAGAGCGAAGCGAACAATTTGATTTT 5240  
Db 4915 AGGCAAGGCTACTAAAACTGAGGGGAGTGAGAGCGAAGCGAACAATTTGATTTT 4974  
QY 5241 ATTTCTATCTTTTAGTGCATAGAGTATCTTATTTGCTCTATAAACTATTTAGCAG 5300  
Db 4975 ATTTCTATCTTTTAGTGCATAGAGTATCTTATTTGCTCTATAAACTATTTAGCAG 5034  
QY 5301 CATATAGATTTTAAATAGGTTCATTAAAGTTGACATATTAGAGAGGAAATCTTGG 5360  
Db 5035 CATATAGATTTTAAATAGGTTCATTAAAGTTGACATATTAGAGAGGAAATCTTGG 5094  
QY 5361 AGAAATATTTGAAGAACCCGATTTACATGGATTGGATTAGTTCTTGTGGTTACGTTG 5420  
Db 5095 AGAAATATTTGAAGAACCCGATTTACATGGATTGGATTAGTTCTTGTGGTTACGTTG 5154  
QY 5421 TAATAAAGTGTGATTTTTCATTTTGGTGTGCTGCTGTTGTTAGTATTTGCTA 5480  
Db 5155 TAATAAAGTGTGATTTTTCATTTTGGTGTGCTGCTGTTGTTAGTATTTGCTA 5214  
QY 5481 GTCAAGTGTATTAATA 5497  
Db 5215 GTCAAGTGTATTAATA 5231

RESULT 8  
AF038666

LOCUS AF038666 8072 bp DNA circular SYN 24-APR-1998  
DEFINITION Cloning vector pFUN, complete sequence.  
ACCESSION AF038666  
VERSION AF038666.1 GI:3043924  
KEYWORDS  
SOURCE  
ORGANISM  
Cloning vector pFUN  
Cloning vector pFUN  
artificial sequences; vectors.  
REFERENCE  
1 (bases 1 to 8072)  
Poquet, I., Ehrlich, S.D. and Gruss, A.  
An export-specific reporter designed for gram-positive bacteria:  
application to *Lactococcus lactis*  
J. Bacteriol. 180 (7), 1904-1912 (1998)  
JOURNAL  
MEDLINE  
98196737  
PUBMED  
9537391  
REFERENCE  
2 (bases 1 to 8072)  
Poquet, I. and Gruss, A.  
Direct Substitution  
Submitted (15-DEC-1997) Laboratoire de Genetique Appliquee-URLCA,  
Institut National de la Recherche Agronomique, CRJ, Jouy en Josas  
78352, France  
FEATURES  
Location/Qualifiers  
source  
1..8072  
/organism="Cloning vector pFUN"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:74538"  
/note="plasmid vector that creates translational fusions  
with a nuclelease reporter without a signal peptide"  
misc\_feature  
1..4580  
/note="p1252 moiety; similar to sequence deposited under  
GenBank Accession Number AF039139"  
misc\_feature  
4581..7472  
/note="similar to BlueScript"  
terminator  
7473..7507  
/note="trpA terminator"  
misc\_feature  
7512..7547  
/note="multiple cloning site"  
gene  
7543..8010  
/gene="delta-nuc"  
/note="truncated nuc gene potentially encoding deltaSP-Nuc  
ORF"  
CDS  
7543..8010  
/gene="delta-nuc"  
/note="nuclease from *Staphylococcus aureus* devoid of its  
signal peptide and the thirteen amino acids at the  
N-terminus of mature protein"  
/transl\_table=1  
/product="deltaSP-Nuc"  
/protein\_id="AAC14660.1"  
/cds\_xref="GI:3043925"  
/translation="DPTVTSATSKLHKSPALIKADGTVKLMYKQPMPTFRLL  
VDTEPKPKGVEKYGEASATKQVNAKKIEVEFDKQRTDKYGRGLAYADG  
KWNVEALVRQGLAKVAVYKPNNTHECHLRKSEAQAQKKEKLNWSEDNADSGQ"  
ORIGIN  
Query Match 73.2%; Score 4025.2; DB 12; Length 8072;  
Best Local Similarity 91.4%; Fred. No. 0;  
Matches 4422; Conservative 0; Mismatches 153; Indels 265; Gaps 4;  
QY 658 CTAGAATCGATACGATTTTGAAGTGGCAACAGATATAAAGAGCAGTTTAAATTTGTTGC 717  
Db 1 CTAGAATCGATACGATTTTGAAGTGGCAACAGATATAAAGAGCAGTTTAAATTTGTTGC 60  
QY 718 TGAATCTTTAAACAGCAATATCAATCTGTGCAACAGATAGCGACAGAGGCGA 777  
Db 61 TGAATCTTTAAACAGCAATATCAATCTGTGCAACAGATAGCGACAGAGGCGA 120  
QY 778 AAACATTCCTGGTCGATCATTTATAAGCAAAATGCCTTTTCTTAAAGATAAAGCTATAA 837  
Db 121 AAACATTCCTGGTCGATCATTTATAAGCAAAATGCCTTTTCTTAAAGATAAAGCTATAA 180  
QY 838 AAGACTATGATCAATAGTTTAGAAAAAGATGTGATCGTAGCGGTTTTTCAAAATTTGCA 897

181	Db		AAGACTATGGATCAATAGTTT	AGAAAAAGATGTCATCCGTAGCGGTTT	CAAAAATTGCA	240
898	Qy		ACCGGAATGAATTACTATCCCTTTTAT	CAAGAAAGCGCAAAAGAAAAACGAAATGATACA	957	
241	Db		ACCGGAATGAATTACTATCCCTTTTAT	CAAGAAAGCGCAAAAGAAAAACGAAATGATACA	300	
958	Qy		CCAATCAGTGCAAAAAAGATATAAT	TGGGAGATGAAGCGTTTCGTGTTTCGTGCTGACTTG	1017	
301	Db		CCAATCAGTGCAAAAAAGATATAAT	TGGGAGATGAAGCGTTTCGTGTTTCGTGCTGACTTG	360	
1018	Qy		CACCATATCATAAAAATCGAAACAGCAAGAAAT	TGGCGGAACGTAAAGAAAGTTATGGAA	1077	
361	Db		CACCATATCATAAAAATCGAAACAGCAAGAAAT	TGGCGGAACGTAAAGAAAGTTATGGAA	420	
1078	Qy		ATAAGACITTAGAAGCAAACTTTAAGAGTGTGTT	TGATAGTCAGTATCTTAAAAATTTTGTAT	1137	
421	Db		ATAAGACITTAGAAGCAAACTTTAAGAGTGTGTT	TGATAGTCAGTATCTTAAAAATTTTGTAT	480	
1138	Qy		AATAGGAATTGAAGTTAAATTAGATGCTAAAAAT	TGTAATTAAGAAGGAGTGAATACAT	1197	
481	Db		AATAGGAATTGAAGTTAAATTAGATGCTAAAAAT	TGTAATTAAGAAGGAGTGAATACAT	540	
1198	Qy		GAACAAAAATATAAATATCTCAAAACTTTTAA	CGAGTGAAAAAGTACTCAACCAAAAT	1257	
541	Db		GAACAAAAATATAAATATCTCAAAACTTTTAA	CGAGTGAAAAAGTACTCAACCAAAAT	600	
1258	Qy		AATAAAACAAATGAAATTTAAAGAAACGGATAC	CGTTTACGAAATTTGGAAACAGGTAAAGG	1317	
601	Db		AATAAAACAAATGAAATTTAAAGAAACGGATAC	CGTTTACGAAATTTGGAAACAGGTAAAGG	660	
1318	Qy		GCATTTACAGCGAACTGGCTTAAATAAGTAAC	AGTAAACGTCATTCGAATTTAGACAG	1377	
661	Db		GCATTTACAGCGAACTGGCTTAAATAAGTAAC	AGTAAACGTCATTCGAATTTAGACAG	720	
1378	Qy		TCATCTATTCAACTTATCGTCAGAAAAATTA	AAAACTGAATCTCGTCTCACTTAATTTCA	1437	
721	Db		TCATCTATTCAACTTATCGTCAGAAAAATTA	AAAACTGAATCTCGTCTCACTTAATTTCA	780	
1438	Qy		CCAGATATTCTACAGTTTCAATTTCCCTTA	CAAAACAGAGGTATAAAATTTCTGGGAGTAT	1497	
781	Db		CCAGATATTCTACAGTTTCAATTTCCCTTA	CAAAACAGAGGTATAAAATTTCTGGGAGTAT	840	
1498	Qy		TCCTTTACCATTTAAGCACACAAAATTA	TAAAAAAGTGGTTTTTGAAGCCATCGTCTGA	1557	
841	Db		TCCTTTACCATTTAAGCACACAAAATTA	TAAAAAAGTGGTTTTTGAAGCCATCGTCTGA	900	
1558	Qy		CATCTATCTGATTTGTTGAAGAGGATTTCTA	CAAGCGTACCTTGGAATTTACCCGAACACT	1617	
901	Db		CATCTATCTGATTTGTTGAAGAGGATTTCTA	CAAGCGTACCTTGGAATTTACCCGAACACT	960	
1618	Qy		AGGGTTGCTCTTTGCACACTCAAGTCTCGAT	TGCAATTTGCTTAAGCTGCCAGCGGAATG	1677	
961	Db		AGGGTTGCTCTTTGCACACTCAAGTCTCGAT	TGCAATTTGCTTAAGCTGCCAGCGGAATG	1020	
1678	Qy		CTTTTCATCCTTAAACCAAAAGTAAACAGT	GTCTTAATAAACTTACC CGCCATACCAGA	1737	
1021	Db		CTTTTCATCCTTAAACCAAAAGTAAACAGT	GTCTTAATAAACTTACC CGCCATACCAGA	1080	
1738	Qy		TGTTCCAGATAAATAATTGGAAGCTATATAC	GTACTTTGTTTCAAAATGGGTCAATCGAGA	1797	
1081	Db		TGTTCCAGATAAATAATTGGAAGCTATATAC	GTACTTTGTTTCAAAATGGGTCAATCGAGA	1140	
1798	Qy		ATATCGTCAACTGTTTACTAAAAATCAGTTT	TCATCAAGCAATGAAACACGCCAAAGTAAA	1857	
1141	Db		ATATCGTCAACTGTTTACTAAAAATCAGTTT	TCATCAAGCAATGAAACACGCCAAAGTAAA	1200	
1858	Qy		CAATTTAAGTACCGTTACTTATGAGCAAGTAT	TGCTATTTTTTAATAGTTATCTATTAAT	1917	
1201	Db		CAATTTAAGTACCGTTACTTATGAGCAAGTAT	TGCTATTTTTTAATAGTTATCTATTAAT	1260	
1918	Qy		TAAACGGGAGGAAATAATTCTATGATCGCT	TTTCTGTAATTTGGAAAGTTACACGTTACTA	1977	

D	b	1261	TACGGGAGGAATAATCTTATAGTCGCCTTTTGTAATTTGGAAAGTTACACGTACTA	1320
Q	y	1978	AAGGGAATTAGATAAAATTTATTAGGTATCTACTGCAGCCTTCCAAGGAGCTAAAAGAGGT	2037
D	b	1321	AAGGGAATTAGATAAAATTTATTAGGTATCTACTGCAGCCTTCCAAGGAGCTAAAAGAGGT	1380
Q	y	2038	CCCTAGCGCTTTATCATGCGGAGCTCGGATCATATGCAAGACAAAAATAAACCTCGCAC	2097
D	b	1381	CCCTAGCGCT-----	1390
Q	y	2098	AGCACTTGGAGAAATGGACGAATCGAGAAAAACCCTCTTTTACGCTGGATTACATATCTAA	2157
D	b	1391	-----TAGAATCGCTTTAGGAACAACGATCCAGTCCA	1422
Q	y	2158	TAAAGCGGTAAAGGACGCGGTTCAAAGAAGTTTAAATAAAGGAGAGACAAATCAATCGATT	2217
D	b	1423	ATAATCGTCGATAAAAACTTTGAAAAGGTTGGTGAATTAACCTACTCTTTGGAAATAATC	1482
Q	y	2218	AGCTAGAACTATATTTTTTCGACAAACGTGCAGAAATTTAGAGAAAGCTCTCTCCAAGACCA	2277
D	b	1483	ACAAATCACAAGTGATTAATCAAAATCACAAGTGATTAATCACTTGTTTATTAG--	1538
Q	y	2278	GTTACAAAGAGCTAGTGCCTAAACATAATTTATTAAAGCTATTAAGTGTGTGGAAACACTGT	2337
D	b	1539	-----ATATAAAGACTATTAATTTAAATAAAGCGTGA	1570
Q	y	2338	ATATATGGNAAGCCCTAGAGAAATTAAGAACAAGAGGAGAAATTTAGACAAAGATTAAAT	2397
D	b	1571	ATTTTATTACAAAAAGAGGGGGGAAACTCTTGGAACTAGCATTTTAGAGAAAGCTTAA--	1628
Q	y	2398	GCCATATGCGTGGCGGTTAGGATGGGAACATATCAATTTTCTTTGGAGAAATACAAATTTGA	2457
D	b	1629	-----AAAAAGATGA	1638
Q	y	2458	AGGATTAATGACACACTGGGCAATTAATTTACGTCCTTTTACGTTATAAAGAGCGGTTTTA	2517
D	b	1639	GAGGTACCAATCAAAAGNAAAAAT-----	1662
Q	y	2518	TTCTTAATAAAGCGCTCTTTTTATAGAAAAAATCCTTAGCGTGGTTTTTTCCGAAATG	2577
D	b	1663	-----	1662
Q	y	2578	CTGCGGTAACCCAGAATTTAGAAATGAGTAGATCAAAATTTATTCACGAATAGAAATCAGGA	2637
D	b	1663	-----TCTCCCAAGAAATTAGAAATGAGTAGATCAAAATTTATTCAGAAATAGAAATCAGGA	1715
Q	y	2638	AAATCAGATCCAACTATAAAACACTAGAACAAATTCGCAAGTTTAACCTACTCAACGCTA	2697
D	b	1716	AAATCAGATCCAACTATAAAACACTAGAACAAATTCGCAAGTTTAACCTACTCAACGCTA	1775
Q	y	2698	GTAGTGGATTAAATCCCAATGAGCCCAACAGAACCCAGAGCCGAGAAACAGAAATCAGAACAA	2757
D	b	1776	GTAGTGATTTAAATCCCAATGAGCCCAACAGAACCCAGAGAAACAGAAATCAGAACAA	1835
Q	y	2758	GTACATTTGATTTAGAAATGGAGAGAGAAAGCAATCAGCTCGTGTGCAATTAATGCCAC	2817
D	b	1836	GTAACATTTGATTTAGAAATGGAGAGAGAAAGCAATCAGCTCGTGTGCAATTAATGCCAC	1895
Q	y	2818	GAATTCGTTGCTTATTTTTTTTTTAAAAAGCGGTATCTAGATATAACGAAACAAACGAACTG	2877
D	b	1896	GAATTCGTTGCTTATTTTTTTTTTAAAAAGCGGTATCTAGATATAACGAAACAAACGAACTG	1955
Q	y	2878	AATAGAAACGMAAAAGACGACATTTATTAATTTGTTTGCAGCACTTTTATAAAT	2937
D	b	1956	AATAGAAACGMAAAAGACGACATTTATTAATTTGTTTGCAGCACTTTTATAAAT	2015
Q	y	2938	GATAGCCCGATTAAGATTGCAACCAACGCTTATCAGTTTGTAGTGAATGAATCTTCCCCT	2997
D	b	2016	GATAGCCCGATTAAGATTGCAACCAACGCTTATCAGTTTGTAGTGAATGAATCTTCCCCT	2075
Q	y	2998	CGTAAGAGTTATTTAAATTAATTTGTTTGCAGACGCTATATAAACCGTACTATCATATA	3057
D	b	2076	CGTAAGAGTTATTTAAATTAATTTGTTTGCAGACGCTATATAAACCGTACTATCATATA	2135









Db 241 GTGTTGATAGTGCAGTATCTTAAAAATTTTGTATATAATAGGAATGAAAGTTAAATTTAGATGC 300  
Qy 1165 TAAAAATTTGTAATTAAGAGGAGTGAATACATGAACAAAAATATAAAAAATATCTCAAAA 1224  
Db 301 TAAAAATTTGTAATTAAGAGGAGTGAATACATGAACAAAAATATAAAAAATATCTCAAAA 360  
Qy 1225 CTTTTTAAACGAGTGAAGAGTACTCAACCAATATAAACAATTCGAATTTAAAGAAAC 1284  
Db 361 CTTTTTAAACGAGTGAAGAGTACTCAACCAATATAAACAATTCGAATTTAAAGAAAC 420  
Qy 1285 CGATACCGTTTACGAAATTTGAAACAGCTAAAGGCAATTTAACGACGAAACTGGCTAAAAAT 1344  
Db 421 CGATACCGTTTACGAAATTTGAAACAGCTAAAGGCAATTTAACGACGAAACTGGCTAAAAAT 480  
Qy 1345 AAGTAAACAGGTAACGCTATTGAAATTAAGAGTGAATACATGAACAAAAATATAAAAAATATCTCAAAA 1404  
Db 481 AAGTAAACAGGTAACGCTATTGAAATTAAGAGTGAATACATGAACAAAAATATAAAAAATATCTCAAAA 540  
Qy 1405 ATTAAAACTGAATCTCGTGTCACTTTAATTTACCAAGATATTTCTACAGTTTCAATTTCC 1464  
Db 541 ATTAAAACTGAATCTCGTGTCACTTTAATTTACCAAGATATTTCTACAGTTTCAATTTCC 600  
Qy 1465 TAAACAAACAGAGTATAAAATTTTGGAGTATTCCTTACCAATTTAAGCAACAATTTAT 1524  
Db 601 TAAACAAACAGAGTATAAAATTTTGGAGTATTCCTTACCAATTTAAGCAACAATTTAT 660  
Qy 1525 TAAAGAGTGTGTTTTGAAGCCATCGCTGTGACATCTATCTGATTTGGAAGAGGAT 1584  
Db 661 TAAAGAGTGTGTTTTGAAGCCATCGCTGTGACATCTATCTGATTTGGAAGAGGAT 720  
Qy 1585 CTAAAGCGTACCTTGGATATTACCGAACAACATAGGTTGCTCTTGCACACATCAAGTCTC 1644  
Db 721 CTAAAGCGTACCTTGGATATTACCGAACAACATAGGTTGCTCTTGCACACATCAAGTCTC 780  
Qy 1645 GATTACGAAATTCGTTAAGCTGCAGCGGAAATGCTTTTCATCTAAACCAAGTAACAG 1704  
Db 781 GATTACGAAATTCGTTAAGCTGCAGCGGAAATGCTTTTCATCTAAACCAAGTAACAG 840  
Qy 1705 TGTCTTAATAAACTTACCGCCATACCAAGATGTTCCAGATAAATATTGGAAGCTATA 1764  
Db 841 TGTCTTAATAAACTTACCGCCATACCAAGATGTTCCAGATAAATATTGGAAGCTATA 900  
Qy 1765 TAGCTACTTTGTTTCAAAATGGGTCAATCGAGAAATATCGTCAACTGTTTACTAAAAATCA 1824  
Db 901 TAGCTACTTTGTTTCAAAATGGGTCAATCGAGAAATATCGTCAACTGTTTACTAAAAATCA 960  
Qy 1825 GTTTCATCAGCAATGAACACGCCAAAGTAACAAATTTAAGTACCGTTACTTTATGAGCA 1884  
Db 961 GTTTCATCAGCAATGAACACGCCAAAGTAACAAATTTAAGTACCGTTACTTTATGAGCA 1020  
Qy 1885 AGTATTGCTATTTTTAAATAGTTATCTATTATTAAACGGGAGAAATAATTTCTATGAGTC 1944  
Db 1021 AGTATTGCTATTTTTAAATAGTTATCTATTATTAAACGGGAGAAATAATTTCTATGAGTC 1080  
Qy 1945 GCTTTTGTAAATTTGGAAGTTACAGTTTACTAAAGGGATGTAGATTAATTTAGGTA 2004  
Db 1081 GCTTTTGTAAATTTGGAAGTTACAGTTTACTAAAGGGATGTAGATTAATTTAGGTA 1140  
Qy 2005 TACTACTGACAGCTTCCAAAGGAGCTAAAGAGTCCCTTAGCGCTCTTTATCATGGGGAAGCT 2064  
Db 1141 TACTACTGACAGCTTCCAAAGGAGCTAAAGAGTCCCTTAGCGCTCTTTATCATGGGGAAGCT 1184  
Qy 2065 CGGATCATATGCAAGACAAATAAATCTCGCAACAGCACTTGGAGAAATGGGAGCAATCGA 2124  
Db 1185 ----- 1184  
Qy 2125 GAAAAACCTCTTTACGCTGGATTACATATCTAATAAGCCGTAAAGGAGCGGTTCAAAA 2184  
Db 1185 --AGATCGCTTTAGGAAACACGATCCGCTCCATATATCTCGATTAATAAATCTTTGAAA 1242  
Qy 2185 AGGTTTAAATTAAGGAGAGCAATCAATGCAATTTAGCTAGCAATATATTTTTCGACACG 2244  
Db 1243 AGGTTTGGAAATTAACCTACTTTTGGAAATAATCACAAATCACAAATGATTAATTAAT 1302

Qy 2245 TGGAGAAATTTAGAGAACGTGCTCTCCAAAGACCAGCTTACAAAGAGCTAGTGCACCTAAACAT 2304  
Db 1303 CACAAGTGAATTAATCAGCTGTTGTTATTAG----- 1331  
Qy 2305 AATTATTAAACGCTATAAGTGTGGAACACACTGTATATATGAAAAAGCCGTAGAGAAT 2364  
Db 1332 -ATATTAAAGCTATATTTAAATTAAGCGTGAATTTTATTACACAAAGAGAGGGGAG 1390  
Qy 2365 AAAAGCAAGAGAGAAATTTAGAGAAATTTAATGCCATATCGGTGGCGGTTAGGATGGA 2424  
Db 1391 AAACCTTGGAACTAGCAATTTAGAGAAAGCTTA----- 1421  
Qy 2425 ACATATCAATTTTCTTGAGAAATACAAATTTGAAGGATTACATGACACTGGGCAAAATGA 2484  
Db 1422 -----AAAAGATGAGAGGTACCAATCAAAGAAAAAT--- 1455  
Qy 2485 TTTAGTCTTTTACGTATAAAGAGCGGCTTTTATTCTTAATATAACGGCTCTTTTATTAG 2544  
Db 1456 ----- 1455  
Qy 2545 AAAAAATCCTTAGCGTGGTTTTTTCGAAATGCTGGCGGTACCCCAAGRAATTTAGAAATG 2604  
Db 1456 -----TCTCCCAAGAAATTTAGAAATG 1475  
Qy 2605 AGTAGATCAAAATTTTACGAATAGAAATCAGGAAAAATCAGATCCAAACCATAAAAACCTA 2664  
Db 1476 AGTAGATCAAAATTTTACGAATAGAAATCAGGAAAAATCAGATCCAAACCATAAAAACCTA 1535  
Qy 2665 GAAACAAATTTGAAAGTTAACTCAACGCTAGTGTGATTTTAAATCCCAAAATGAGCCA 2724  
Db 1536 GAAACAAATTTGAAAGTTAACTCAACGCTAGTGTGATTTTAAATCCCAAAATGAGCCA 1595  
Qy 2725 ACAGAACAGAGCCAGAAACAGAAATCAGAAACAGTAAATTTAGAAATGGAAGAA 2784  
Db 1596 ACAGAACAGAGCCAGAAACAGAAATCAGAAACAGTAAATTTAGAAATGGAAGAA 1655  
Qy 2785 GAAAAAGCAATGACTTCGTGTGAATTAATGACGAAATCGTGTGTTATTTTTTTTAAAA 2844  
Db 1656 GAAAAAGCAATGACTTCGTGTGAATTAATGACGAAATCGTGTGTTATTTTTTTTAAAA 1715  
Qy 2845 GCGGTACTAGATATAACGAAACAAAGAACTGAATGAAACGAAAGAACGACATGACA 2904  
Db 1716 GCGGTACTAGATATAACGAAACAAAGAACTGAATGAAACGAAAGAACGACATGACA 1775  
Qy 2905 CATTTTATAAATGTTTGAACGACATTTTATAAATGCATAGCCGATAGATTTGCCAAACCA 2964  
Db 1776 CATTTTATAAATGTTTGAACGACATTTTATAAATGCATAGCCGATAGATTTGCCAAACCA 1835  
Qy 2965 AGCTTATCAGTTAGTCAAGATGAATCTTCCCTCGTAAGAGGTTATTTAATTAATTTGT 3024  
Db 1836 AGCTTATCAGTTAGTCAAGATGAATCTTCCCTCGTAAGAGGTTATTTAATTAATTTGT 1895  
Qy 3025 TTGAAGACGGTATATAACCGTACTATCATTATATAGGGAATCAGAGATTTTCAAGTAT 3084  
Db 1896 TTGAAGACGGTATATAACCGTACTATCATTATATAGGGAATCAGAGATTTTCAAGTAT 1955  
Qy 3085 CTAAGCTACTGAATTTAAGAAATGTTAAGCAATCAATCGGAAATCGTTTGTGATTTGCTTTT 3144  
Db 1956 CTAAGCTACTGAATTTAAGAAATGTTAAGCAATCAATCGGAAATCGTTTGTGATTTGCTTTT 2015  
Qy 3145 TTGTATTCAATTTATAGAGGTGGAGTTTGTATGATCATGATGAATGTAAACCTTATATA 3204  
Db 2016 TTGTATTCAATTTATAGAGGTGGAGTTTGTATGATCATGATGAATGTAAACCTTATATA 2075  
Qy 3205 AAAAAATGTTTATGGAGATAGAAATTTAGCAATATCTATACACTAGAAACGTTTAAAG 3264  
Db 2076 AAAAAATGTTTATGGAGATAGAAATTTAGCAATATCTATACACTAGAAACGTTTAAAG 2135  
Qy 3265 AAAGAGTTAGAAAGAGAAATCTACTTTAGAAACAAATCAGATAAGTATTTTCTTCG 3324  
Db 2136 AAAGAGTTAGAAAGAGAAATCTACTTTAGAAACAAATCAGATAAGTATTTTCTTCG 2195

Qy	3325	GAGGGGAAGATTATATATATAAGTTTAATAGAAAATAACAAAATAATTTTATTCGATTAGT	3384
Db	2196	GAGGGGAAGATTATATATATAAGTTTAATAGAAAATAACAAAATAATTTTATTCGATTAGT	2255
Qy	3385	GGAAAAAATTGACTTATAAGGAAAAAATCTTTTCAAACAATGCAATATTGAACAACG	3444
Db	2256	GGAAAAAATTGACTTATAAGGAAAAAATCTTTTCAAACAATGCAATATTGAACAACG	2315
Qy	3445	TTGAATGAAAAAGCAAAACCAAGTTAAATTAAACAACCTTATTATAGGATTATAGGAAAG	3504
Db	2316	TTGAATGAAAAAGCAAAACCAAGTTAAATTAAACAACCTTATTATAGGATTATAGGAAAG	2375
Qy	3505	GAGAACGCTGAATGAATATCCCTTTTGTGTAGAACTCGCTTCATGACGGCTGTGTTA	3564
Db	2376	GAGAACGCTGAATGAATATCCCTTTTGTGTAGAACTCGCTTCATGACGGCTGTGTTA	2435
Qy	3565	AAGTACAAAATTTAAAAATAGCTAAAAATTCGGCTCAATCACTACCAAGCGAGTAAAAACAAA	3624
Db	2436	AAGTACAAAATTTAAAAATAGCTAAAAATTCGGCTCAATCACTACCAAGCGAGTAAAAACAAA	2495
Qy	3625	GGGGCTATTTTGGGTATCGCTCAAAATCAAGCATGATTCGGGGTCTGTGTGTTGTTCTG	3684
Db	2496	GGGGCTATTTTGGGTATCGCTCAAAATCAAGCATGATTCGGGGTCTGTGTGTTGTTCTG	2555
Qy	3685	ACTTCCGAGGAAGCGATTCAAGAAAAATCAAGATACATTTACATTTGGACACCAACGTT	3744
Db	2556	ACTTCCGAGGAAGCGATTCAAGAAAAATCAAGATACATTTACATTTGGACACCAACGTT	2615
Qy	3745	TATCGTTATGGAAAGCTATGACAGCGAAAAACGGTTCATACAGAAAGCATTCCTGAAAAC	3804
Db	2616	TATCGTTATGGAAAGCTATGACAGCGAAAAACGGTTCATACAGAAAGCATTCCTGAAAAC	2675
Qy	3805	AATTTAAGACAAATCAATACCTCTTTTATTGATTTTGATATTCACACGGCAAAAGAACT	3864
Db	2676	AATTTAAGACAAATCAATACCTCTTTTATTGATTTTGATATTCACACGGCAAAAGAACT	2735
Qy	3865	ATTTACAGAAAGCGATTTTTTAAACAAACGGCTATTGATTTAGGTTTTATGCTACTATGATT	3924
Db	2736	ATTTACAGAAAGCGATTTTTTAAACAAACGGCTATTGATTTAGGTTTTATGCTACTATGATT	2795
Qy	3925	ATCAATCTCATAAAGTTATCAAGCATATTTTGTTTTAGAAACGCCAGTCTATGTGACT	3984
Db	2796	ATCAATCTCATAAAGTTATCAAGCATATTTTGTTTTAGAAACGCCAGTCTATGTGACT	2855
Qy	3985	TCAAAATCAGAATTTAAATCTGTCAAAGCAGCAAAATATTTTCGAAAAATATCCGAGAA	4044
Db	2856	TCAAAATCAGAATTTAAATCTGTCAAAGCAGCAAAATATTTTCGAAAAATATCCGAGAA	2915
Qy	4045	TATTTTGGAAAGTCTTTGGCAGTTGATCTAAACGTGTAAATCATTTTGGTATTTGCTCGCAT	4104
Db	2916	TATTTTGGAAAGTCTTTGGCAGTTGATCTAAACGTGTAAATCATTTTGGTATTTGCTCGCAT	2975
Qy	4105	CCAAGAACGCAATGTAGAAATTTTTTGATCTCTAAATACCGTTTATCTCTTCAAAGAAATGG	4164
Db	2976	CCAAGAACGCAATGTAGAAATTTTTTGATCTCTAAATACCGTTTATCTCTTCAAAGAAATGG	3035
Qy	4165	CAAGATTGGTCTTTCAAACAAAACAGATAATAAAGGGCTTACTCGTCAAGTCTAACCGTT	4224
Db	3036	CAAGATTGGTCTTTCAAACAAAACAGATAATAAAGGGCTTACTCGTCAAGTCTAACCGTT	3095
Qy	4225	TTAAGCGGTACAGAGCAAAAAACAGTAGATGAACCCCTGGTTTAACTCTCTTTATTGCAAC	4284
Db	3096	TTAAGCGGTACAGAGCAAAAAACAGTAGATGAACCCCTGGTTTAACTCTCTTTATTGCAAC	3155
Qy	4285	GAAACGAAATTTTTCAGGAGAAAAGGGTTTAAATAGGCGTAAATACGCTCATGTTTACCCCTC	4344
Db	3156	GAAACGAAATTTTTCAGGAGAAAAGGGTTTAAATAGGCGTAAATACGCTCATGTTTACCCCTC	3215
Qy	4345	TCTTTTACGCTACTTTTGTTCAGGCTATTCATTCGAAACGGTCCGAAATATATATATGTTTGAG	4404
Db	3216	TCCTTTTACGCTACTTTTGTTCAGGCTATTCATTCGAAACGGTCCGAAATATATATATGTTTGAG	3275
Qy	4405	TTTTAAATAATCGATTAGATCAACCCCTTTAGAGAAAAAAGAAAGTATCAAAATTTGTAGAAGT	4464

[illegible]

[illegible]

```

/codon_start=1
/transl_table=11
/product="Repe"
/protein_id="AAC38603.1"
/db_xref="GI:3023047"
/translation="MNIPFVETVLHGLLKYFKFKSIRSIITPKGSKGALFAYRS
KSSMIGRGVLTSEALQENQDPFHTWPNVRYGYADENPSYKPGHSENNLRIN
TFDFDHTAKETISASDILTAIDLGFMPTMIKSDKGQYFVLETPVYTSKSE
FKSVAKIISQNIREFVGLSLPVDLTGHNHFGIARIPRTDNVEFFDPNRYSKWOD
WSFQTDNKGTRSLVLSGTEGKQVDEFEWLLHETKFSGEKGLIGNNWMTL
SLAYFSSGYSLETCEYNNFENRLDQPLEKEVIKIVRSAYSENYOGANREYITLIC
KAWYSOLTSKDLFVROGWFKPKKRGORVHLSEWKEDLMAYISEKSDVYKPLVT
TKSIREVLIGPETHLDKLLKVKANOEIFPKIPKGRNGGIQLASVKSLLLSIIKVK
EEKESYIKALNTSFDLEHTFIQETFLNKLSRPKTDQTLDFSYDTG"
rep_origin
D-loop
  4606..4648
  4631..4864
  /note="major D-loop forms observed with res+ top-
  plasmids"
  4631..4827
  /note="major D-loop forms observed with res- top-
  plasmids"
  4631
  /note="initiation site of leading strand synthesis"
  4648..7906
  /note="secondary replication functions"
  4712..4856
  /note="primosome assembly site ssIIA"
  4852..7906
  /note="segregational stability functions"
  4852..4951
  /note="resolution site res"
  /bound_moiety="resolvase"
  4861..4870
  /note="site of Res-beta strand exchange during resolution"
  4863..4865
  /note="main DNA polymerase arrest sites mapped in vivo
  with res- top+ plasmids"
  4942..4948
  /note="orfG"
  4954..5091
  /function="unknown"
  /note="orfG; putative"
  /codon_start=1
  /transl_table=11
  /protein_id="AAC38604.1"
  /db_xref="GI:3023048"
  /translation="MEKYLKNPITWIGLVLVWTLTKSSEFLIFGVCVLLLVFASQS
  D"
  5145..5762
  /gene="res_beta"
  5145..5762
  /gene="res_beta"
  /function="resolution of plasmid multimers and control of
  DNA polymerase progression"
  /codon_start=1
  /transl_table=11
  /product="resolvase"
  /protein_id="AAC38605.1"
  /db_xref="GI:3023049"
  52.9%; Score 2908.4; DB 1; Length 8199;
  Best Local Similarity 100.08; Pred. No. 0;
  Matches 2509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
  QY 2588 CCCAAGATTAGAAATGATGATCAATATTTCAGGAATAGAAATCAGGAAATCAGATC 2647
  DB 2185 CCCAAGATTAGAAATGATGATGATCAATATTTCAGGAATAGAAATCAGGAAATCAGATC 2244
  QY 2648 CAACCATAAACACTAGAACAAATTCGAAAGTTAACTCAACCGCTAGTAGTGATT 2707
  DB 2245 CAACCATAAACACTAGAACAAATTCGAAAGTTAACTCAACCGCTAGTAGTGATT 2304
  QY 2708 TAATCCCAATGAGCCACAGAACCCAGAGCCGAGAACAGAAATCAGAAATCAGAAATCAGTTC 2767

```

```

  2305 TAATCCCAATGAGCCACAGAACCCAGAACGAAACAGAAATCAGAAATCAGTTCAGTTCG 2364
  QY 2768 ATTTAGAAATGGAAGAGAGAGAAACCAATGACTTCGTGTGAATTAATGACGAAATCGTTG 2827
  DB 2365 ATTTAGAAATGGAAGAGAGAGAGAAACCAATGACTTCGTGTGAATTAATGACGAAATCGTTG 2424
  QY 2828 CTTATTTTAAAAAGCGGTACTACTAGATATACGAAACCAACCACTGATGATGAAACG 2887
  DB 2425 CTTATTTTAAAAAGCGGTACTACTAGATATACGAAACCAACCACTGATGATGAAACG 2484
  QY 2888 AAAAAAGAGCCATGACACATTTTAAAAATGTTTGAACGACATTTTATAAATGATGACCCG 2947
  DB 2485 AAAAAAGAGCCATGACACATTTTAAAAATGTTTGAACGACATTTTATAAATGATGACCCG 2544
  QY 2948 ATAAGATTGCCAAACCAACGCTTATCAGTTAGTACAGTGAATCTTCCCTCGTGAAGAT 3007
  DB 2545 ATAAGATTGCCAAACCAACGCTTATCAGTTAGTACAGTGAATCTTCCCTCGTGAAGAT 2604
  QY 3008 TATTTAAATTAACCTTTGTTTGAAGACGGTATATAACCGTACTATCATTTATATAGGAAATC 3067
  DB 2605 TATTTAAATTAACCTTTGTTTGAAGACGGTATATAACCGTACTATCATTTATATAGGAAATC 2664
  QY 3068 AGAGAGTTTCAAGTATCTAAGCTACTGAAATTAAGAAATGTTTAAGCAATCAATCGGAAA 3127
  DB 2665 AGAGAGTTTCAAGTATCTAAGCTACTGAAATTAAGAAATGTTTAAGCAATCAATCGGAAA 2724
  QY 3128 TCGTTTGATGCTTTTGTGTTTATTAAGAGGTGGAGTTTGTATGATCATGATG 3187
  DB 2725 TCGTTTGATGCTTTTGTGTTTATTAAGAGGTGGAGTTTGTATGATCATGATG 2784
  QY 3188 AATGTAAATCTTATATAAAAAATAGTTTATTTGGAGATAGAAATATAGCAAAATATCTATA 3247
  DB 2785 AATGTAAATCTTATATAAAAAATAGTTTATTTGGAGATAGAAATATAGCAAAATATCTATA 2844
  QY 3248 CACTAGAACGTTTAAAGAGAGTTAGAAAGAGAAATATCTACTTGAACACAAATCAG 3307
  DB 2845 CACTAGAACGTTTAAAGAGAGTTAGAAAGAGAAATATCTACTTGAACACAAATCAG 2904
  QY 3308 ATAAGTATTTTCTTCGGAGGGGAGATTTATATATATAAGTTAATAGAAATTAACAAA 3367
  DB 2905 ATAAGTATTTTCTTCGGAGGGGAGATTTATATATAAGTTAATAGAAATTAACAAA 2964
  QY 3368 TAATTTATTCGATTAGTGAAAAAATTCAGCTTATAAGGAAAAAATCTTTTCAAAAC 3427
  DB 2965 TAATTTATTCGATTAGTGAAAAAATTCAGCTTATAAGGAAAAAATCTTTTCAAAAC 3024
  QY 3428 ATGCAATATTGAAACAGTTGAATGAAAGCAAAACCAAGTTAAATTAACAACTATTTTA 3487
  DB 3025 ATGCAATATTGAAACAGTTGAATGAAAGCAAAACCAAGTTAAATTAACAACTATTTTA 3084
  QY 3488 TAGGATTATAGAAAGAGAACACAGCTGAATGATATCCCTTTTGTGTAGAACTGTGC 3547
  DB 3085 TAGGATTATAGAAAGAGAACACAGCTGAATGATATCCCTTTTGTGTAGAACTGTGC 3144
  QY 3548 TTCATGACGGCTTGTAAAGTCAAAATTTAAAAATAGTAAAAATTCGCTCAATCACTACCA 3607
  DB 3145 TTCATGACGGCTTGTAAAGTCAAAATTTAAAAATAGTAAAAATTCGCTCAATCACTACCA 3204
  QY 3608 AGCCAGGTAAAAAGCAAAAGGGGCTATTTTTCGGTATCGCTCAAAATCAAGCATGATTGCG 3667
  DB 3205 AGCCAGGTAAAAAGCAAAAGGGGCTATTTTTCGGTATCGCTCAAAATCAAGCATGATTGCG 3264
  QY 3668 GTCTGTTGTTGTTCTGACTTCCGAGGAGCGATTCAAGAAATCAAGATACATTTTACAC 3727
  DB 3265 GTCTGTTGTTGTTCTGACTTCCGAGGAGCGATTCAAGAAATCAAGATACATTTTACAC 3324
  QY 3728 ATTTGACACCCCAACGTTTATCGTTTATGAAACGTTATGACAGCAAAACCGTTTATACAGA 3787
  DB 3325 ATTTGACACCCCAACGTTTATCGTTTATGAAACGTTATGACAGCAAAACCGTTTATACAGA 3384
  QY 3788 AGGACATTCGAAACCAATTTAAGACAAATCAATACCTTCTTTTATGATTTGATATTC 3847

```

3385 AAGGACATCTCTGAAACAAATTAAGACAAATCAATCTCTCTTTATTTGATTTGATATTC 3444  
3848 ACACGGCAAAAGAAACTATTTTCAGCAAGCGATATTTTAAACACGGCTATTTGATTTAGGTT 3907  
3445 ACACGGCAAAAGAACTATTTTCAGCAAGCGATATTTTAAACACGGCTATTTGATTTAGGTT 3504  
3308 TTATGCTCTACTGATGATTAATCAATCTGATAAGGTTATCAAGCATATTTTGTTTTAGAAA 3967  
3505 TTATGCTCTACTGATGATTAATCAATCTGATAAGGTTATCAAGCATATTTTGTTTTAGAAA 3564  
3368 CGCCAGCTATGATGATTAATCAATCTGATAAGGTTATCAAGCATATTTTGTTTTAGAAA 4027  
3565 CGCCAGCTATGATGATTAATCAATCTGATAAGGTTATCAAGCATATTTTGTTTTAGAAA 3624  
4028 CGCAAAATATCCGAGATATTTTGGAAAGCTTTTGGCCAGTTGATCTAAACGCTGTAATCAAT 4087  
3625 CGCAAAATATCCGAGATATTTTGGAAAGCTTTTGGCCAGTTGATCTAAACGCTGTAATCAAT 3684  
4088 TTGCTATGCTGATGATTAATCAATCTGATAAGGTTATCAAGCATATTTTGTTTTAGAAA 4147  
3685 TTGCTATGCTGATGATTAATCAATCTGATAAGGTTATCAAGCATATTTTGTTTTAGAAA 3744  
4148 ATTCTTTCAAGATGCGAAGATGCTTTTCAACCAACAGATATTAAGGGCTTTTACTC 4207  
3745 ATTCTTTCAAGATGCGAAGATGCTTTTCAACCAACAGATATTAAGGGCTTTTACTC 3804  
4208 GTTCAAGCTTAACGGTTTAAAGCGGTACAGAGGCAAAACCAAGTAGATGAACCCCTGGT 4267  
3805 GTTCAAGCTTAACGGTTTAAAGCGGTACAGAGGCAAAACCAAGTAGATGAACCCCTGGT 3864  
4268 TTAATCTCTTATGCGAAGATGCTTTTCAACCAACAGATATTAAGGGCTTTTACTC 4327  
3865 TTAATCTCTTATGCGAAGATGCTTTTCAACCAACAGATATTAAGGGCTTTTACTC 3924  
4328 ACCTCATGTTTACCTCTCTTACGCTTACCTTACGCTTACCTTACGCTTACCTTACGCTTACCT 4387  
3925 ACCTCATGTTTACCTCTCTTACGCTTACCTTACGCTTACCTTACGCTTACCTTACGCTTACCT 3984  
4388 AATATAATGTTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4447  
3985 AATATAATGTTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4044  
4448 TCAGAAATGTTAGAGGCTTATTCAGAAACTATCAAGGGCTATTAAGGGCTATTAAGGGCTATTAAG 4507  
4045 TCAGAAATGTTAGAGGCTTATTCAGAAACTATCAAGGGCTATTAAGGGCTATTAAGGGCTATTAAG 4104  
4508 CCATCTCTTCAAGCTTGGGTATCAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4567  
4105 CCATCTCTTCAAGCTTGGGTATCAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4164  
4568 AAGGCTGTTTAAATTCAGAAAGAAAGAGCGAAGCTCAAGCTTCAATTTGCTCAGAT 4627  
4165 AAGGCTGTTTAAATTCAGAAAGAAAGAGCGAAGCTCAAGCTTCAATTTGCTCAGAT 4224  
4628 GGAAGAGAGATTTAATGGCTTATTAATGCGAAGAAAGCGATGATTAACAGCCCTTATTAG 4687  
4225 GGAAGAGAGATTTAATGGCTTATTAATGCGAAGAAAGCGATGATTAACAGCCCTTATTAG 4284  
4688 TGACGACCAAAAGAGATGATGAGAGTGTAGGATTCCTGACGCAATTAAGGCAATTAAGTAAT 4747  
4285 TGACGACCAAAAGAGATGATGAGAGTGTAGGATTCCTGACGCAATTAAGGCAATTAAGTAAT 4344  
4748 TGCTGAAGGTACTGAAGCGCAATCAGGAATTTTCTTTAGATTTAAACAGGAGAAATG 4807  
4345 TGCTGAAGGTACTGAAGCGCAATCAGGAATTTTCTTTAGATTTAAACAGGAGAAATG 4404  
4808 GTGCAATCAACTGCTAGTGTAAATCAATGTTGCTATCGATCAATTAAGTAATTAAGTAAT 4867  
4405 GTGCAATCAACTGCTAGTGTAAATCAATGTTGCTATCGATCAATTAAGTAATTAAGTAAT 4464  
4868 AAGAAAGAGAGCTATTAAGCGCTGACAAATCTTTTGTGCTATGAGTATGAGTATGAGTATGAGTAT 4927  
4465 AAGAAAGAGAGCTATTAAGCGCTGACAAATCTTTTGTGCTATGAGTATGAGTATGAGTATGAGTAT 4524

QY 4928 TTCAAGAGACTTTAAACCAAGCTAGCAGAACCGCTTAAACCGACACACAACTCGATTTGT 4987  
Db 4525 TTCAAGAGACTTTAAACCAAGCTAGCAGAACCGCTTAAACCGACACACAACTCGATTTGT 4584  
QY 4988 TTAGCTATGATACAGGCTGAAATAAAACCGCAGCTATGCCATTAATTAATTAATTAATTAAT 5047  
Db 4585 TTAGCTATGATACAGGCTGAAATAAAACCGCAGCTATGCCATTAATTAATTAATTAATTAAT 4644  
QY 5048 TTAGCTATGATACAGGCTGAAATAAAACCGCAGCTATGCCATTAATTAATTAATTAATTAAT 5107  
Db 4645 TTAGCTATGATACAGGCTGAAATAAAACCGCAGCTATGCCATTAATTAATTAATTAATTAAT 4704  
QY 5108 TTAGCTATGATACAGGCTGAAATAAAACCGCAGCTATGCCATTAATTAATTAATTAATTAAT 5167  
Db 4705 TTAGCTATGATACAGGCTGAAATAAAACCGCAGCTATGCCATTAATTAATTAATTAATTAAT 4764  
QY 5168 ACTGAACAGAGTGAAGGAGAAAGGCTACTAAACCGTCGAGGGGCGAGTGAAGGAGGAGGAGG 5227  
Db 4765 ACTGAACAGAGTGAAGGAGAAAGGCTACTAAACCGTCGAGGGGCGAGTGAAGGAGGAGGAGG 4824  
QY 5228 ACTTCATTTTAAATTTTCTATCTTTTATAGTCAATTAAGTATTAATTAATTAATTAATTAAT 5287  
Db 4825 ACTTCATTTTAAATTTTCTATCTTTTATAGTCAATTAAGTATTAATTAATTAATTAATTAAT 4884  
QY 5288 AACTATTTAGCAGCAATAATAGATTTATTAAGTATTAATTAAGTATTAATTAAGTATTAATTAAG 5347  
Db 4885 AACTATTTAGCAGCAATAATAGATTTATTAAGTATTAATTAAGTATTAATTAAGTATTAATTAAG 4944  
QY 5348 AGGAAATCTTGGAGAAATATTTGAAGAACCGGATTAATTAATTAATTAATTAATTAATTAAT 5407  
Db 4945 AGGAAATCTTGGAGAAATATTTGAAGAACCGGATTAATTAATTAATTAATTAATTAATTAAT 5004  
QY 5408 GTTACGCTGTTTAACTAAAGTAGTGAAATTTTGTGATTTTGTGCTGCTGCTGCTGCTGCTG 5467  
Db 5005 GTTACGCTGTTTAACTAAAGTAGTGAAATTTTGTGATTTTGTGCTGCTGCTGCTGCTGCTG 5064  
QY 5468 TTAGTATTTGCTAGTCAAAAGTATTAATA 5497  
Db 5065 TTAGTATTTGCTAGTCAAAAGTATTAATA 5094

RESULT 11  
AF041239  
LOCUS  
DEFINITION  
Cloning vector pIL253 replication protein D (repD), replication protein E (repE), and adenine methylase (ermAM) genes, complete cds and complete sequence.  
ACCESSION  
AF041239  
VERSION  
AF041239.1 GI:2773391  
KEYWORDS  
Cloning vector pIL253  
SOURCE  
Cloning vector pIL253  
ORGANISM  
artificial sequences; vectors.  
REFERENCE  
1 (bases 1 to 4963)  
AUTHORS  
Simon, D. and Chopin, A.  
TITLE  
Construction of a vector plasmid family and its use for molecular cloning in *Streptococcus lactis*  
JOURNAL  
Biochimie 70 (4); 559-566 (1988)  
MEDLINE  
89000956  
PUBMED  
2844302  
REFERENCE  
2 (bases 1 to 4963)  
AUTHORS  
Chopin, A.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (06-JAN-1998) Genetique Microbienne, INRA, Domaine de Vilvert, Jouy-en-Josas 78352, France  
FEATURES  
source  
1..4963  
/organism="Cloning vector pIL253"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:71279"  
/note="high copy number cloning vector for Gram-positive bacteria; derived from cloning vector pIL252, GenBank





QY 4062 GCCAGTTCATTAAGCGTGAATCAATTTGGTATTCGTCGCATACCAAGACCGCAATGT 4121  
DB 1381 GCCAGTTCATTAAGCGTGAATCAATTTGGTATTCGTCGCATACCAAGACCGCAATGT 1440  
QY 4122 AGAATTTTTCATCTTAATACCGTTATTCCTTTCAAGAATGCGAAGATTGGCTTTCAA 4181  
DB 1441 AGAATTTTTCATCTTAATACCGTTATTCCTTTCAAGAATGCGAAGATTGGCTTTCAA 1500  
QY 4182 ACAAACAGATTAAGGCGTTTACGCTTCAAGTCTAACCGTTTAAAGCGGTACAGAAGG 4241  
DB 1501 ACAAACAGATTAAGGCGTTTACGCTTCAAGTCTAACCGTTTAAAGCGGTACAGAAGG 1560  
QY 4242 CAAAAACAAGTACGACCGCTGTTTAACTCTTATTCACGAAACGAAATTTTCAGG 4301  
DB 1561 CAAAAACAAGTACGACCGCTGTTTAACTCTTATTCACGAAACGAAATTTTCAGG 1620  
QY 4302 AGAAAGGTTTAAATAGGCGTAAATACGTCATGTTTACCGTCTCTTTAGCCTACITTAG 4361  
DB 1621 AGAAAGGTTTAAATAGGCGTAAATACGTCATGTTTACCGTCTCTTTAGCCTACITTAG 1680  
QY 4362 TTCAGGCTATTCATCGAAAGCTCGGAATATAATATGTTTGAAGTTTAAATATCGATPAGA 4421  
DB 1681 TTCAGGCTATTCATCGAAAGCTCGGAATATAATATGTTTGAAGTTTAAATATCGATPAGA 1740  
QY 4422 TCACCCCTTAGAAGAAAGAAAGTAATCAAAATGTTAGAGTGCCTATTTCAGAAACTA 4481  
DB 1741 TCACCCCTTAGAAGAAAGAAAGTAATCAAAATGTTAGAGTGCCTATTTCAGAAACTA 1800  
QY 4482 TCAGGGGCTTAATAGGGAATACATACATCTTTGCAAGCTTGGGTATCAAGTGAATTT 4541  
DB 1801 TCAGGGGCTTAATAGGGAATACATACATCTTTGCAAGCTTGGGTATCAAGTGAATTT 1860  
QY 4542 AACCACTAAAGATTTATTTGTCGCTCAAGGCTGTTTAAATTCAGAAAGAAAGACCGA 4601  
DB 1861 AACCACTAAAGATTTATTTGTCGCTCAAGGCTGTTTAAATTCAGAAAGAAAGACCGA 1920  
QY 4602 ACCTCAACGTTTCATTTGTCAGAAAGAAAGATTTTAAATGCTTATATTCAGCAAAA 4661  
DB 1921 ACCTCAACGTTTCATTTGTCAGAAAGAAAGATTTTAAATGCTTATATTCAGCAAAA 1980  
QY 4662 AAGCGATGATACAGGCTTATTTAGTACGACCAAAAGAGATTTAGAGAGTGTAGG 4721  
DB 1981 AAGCGATGATACAGGCTTATTTAGTACGACCAAAAGAGATTTAGAGAGTGTAGG 2040  
QY 4722 CATTCCTGAACGGACATTAGATAAATGCTGAAGTACTGAAGCGCAATCAGCAAAATTTT 4781  
DB 2041 CATTCCTGAACGGACATTAGATAAATGCTGAAGTACTGAAGCGCAATCAGCAAAATTTT 2100  
QY 4782 CTTTAAAGATTAAACCGAGAAAGAAATGCTGCGCATTCACCTTGTAGTGTAAATCATTTGT 4841  
DB 2101 CTTTAAAGATTAAACCGAGAAAGAAATGCTGCGCATTCACCTTGTAGTGTAAATCATTTGT 2160  
QY 4842 GCTATCGATCATTAAGTAAAAAGAAAGAAAGAAAGCTATATAAGCGCTGACAAA 4901  
DB 2161 GCTATCGATCATTAAGTAAAAAGAAAGAAAGAAAGCTATATAAGCGCTGACAAA 2220  
QY 4902 TCTTTTGCATTAGAGCATATTCATTCAGAGACTTTAAACAAGCTAGCAGAACGCC 4961  
DB 2221 TCTTTTGCATTAGAGCATATTCATTCAGAGACTTTAAACAAGCTAGCAGAACGCC 2280  
QY 4962 TAAACGGACACAACTCGATTTGTTAGCTATGATACAGGCTGAAATTAACCCGCA 5021  
DB 2281 TAAACGGACACAACTCGATTTGTTAGCTATGATACAGGCTGAAATTAACCCGCA 2340  
QY 5022 CTATGCCATTACATTTATATCTATGATACGCTGTTTCTTTGCTGTTTAGCGAAT 5081  
DB 2341 CTATGCCATTACATTTATATCTATGATACGCTGTTTCTTTGCTGTTTAGCGAAT 2400  
QY 5082 GATTAGCAGAAATATACAGAGTAAGATTTTAAATTAATTAATAGGGGGAAGAGAGAGT 5141  
DB 2401 GATTAGCAGAAATATACAGAGTAAGATTTTAAATTAATTAATAGGGGGAAGAGAGAGT 2460  
QY 5142 AGCCCGAAAACTTTTAGTTGGCTTGACTCGAACGAAGTGAAGGAGGCTACTAAAACGT 5201

DB 2461 AGCCCGAAAACTTTTAGTTGGCTTGACTCGAAGTGAAGGAGGCTACTAAAACGT 2520  
QY 5202 CGAGGGGCGAGTGAAGCGGAAGCACTGATTTTAAATTTCTATCTTTTATAGGTC 5261  
DB 2521 CGAGGGGCGAGTGAAGCGGAAGCACTGATTTTAAATTTCTATCTTTTATAGGTC 2580  
QY 5262 ATTAGAGTATACCTATTTCCTCTATAAACTATTAGCAGCATATAAGATTTATTGAAATAG 5321  
DB 2581 ATTAGAGTATACCTATTTCCTCTATAAACTATTAGCAGCATATAAGATTTATTGAAATAG 2640  
QY 5322 GTCATTTAAGTTGAGCATATTAGAGAGGAAATCTTGGAGAAATATTGGAAGAACCCGA 5381  
DB 2641 GTCATTTAAGTTGAGCATATTAGAGAGGAAATCTTGGAGAAATATTGGAAGAACCCGA 2700  
QY 5382 TTACATGGAATTCGATTAGTTCTTGTGCTTACGCTGTTTAACTAAAGTACTGAAATTTT 5441  
DB 2701 TTACATGGAATTCGATTAGTTCTTGTGCTTACGCTGTTTAACTAAAGTACTGAAATTTT 2760  
QY 5442 TGATTTTGGTGTGTGTCTTGTGTTTGTAGTATTTGCTAGTCAAAAGTGAATTAATA 5497  
DB 2761 TGATTTTGGTGTGTGTCTTGTGTTTGTAGTATTTGCTAGTCAAAAGTGAATTAATA 2816

## RESULT 12

AF039139

LOCUS

DEFINITION

Cloning vector pIL252

Cloning vector pIL252

artificial sequences; vectors.

REFERENCE

1 (bases 1 to 4698)

AUTHORS

Simon, D. and Chopin, A.

TITLE

Construction of a vector plasmid family and its use for molecular

cloning in *Streptococcus lactis*

Biochimie 70 (4), 559-566 (1988)

PUBMED

2844302

REFERENCE

2 (bases 1 to 4698)

AUTHORS

Chopin, A.

TITLE

Direct Submission

Submitted (18-DEC-1997) Genetique Microbienne, INRA, Domaine de

Vilvert, Location/Qualifiers

1. 4698

/organism="Cloning vector pIL252"

/mol\_type="genomic DNA"

/db\_xref="taxon:70935"

/lab\_host="Enterococcus faecalis"

/note="low copy number"

494..793

/gene="repD"

494..793

/gene="repD"

/note="repD"

/codon\_start=1

/product="replication protein D"

/protein\_id="AAB94753.1"

/db\_xref="GI:2745826"

/translation="MNEDEKTYIKNSLLSIRKLANIYTLTFPKLEKNIYLETYS

DKYFSSGEDIYIKLIENNKIYISGKLTYSKLSFKSHAILKQLNEKANQVN"

836..2326

/gene="repE"

836..2326

/gene="repE"

/note="repE"

/codon\_start=1



Qy	4362	TTCAAGCTATTCAATCGAAACGTCGGAATATAATATGTTTGGAGTTTAAATATCAATGATAGA	4421	Db	2761	TCATTTTGGTGTGTGCTCTTGTGTAGTATTGCTAGTCAAGAGTAAATA	2816
Qy	4422	TCACCCCTTAGAAGAAAAAGAAAGTAAATCAAAATCTGTAGAAAGTCCTATTTCAGAAAACTA	4481	Db	1741	TCACCCCTTAGAAGAAAAAGAAAGTAAATCAAAATCTGTAGAAAGTCCTATTTCAGAAAACTA	1800
Qy	4482	TCAGGGGCTAATAGGGAATAATACCATCTCTTTGCAAGCTTGGGTATCAAGTGAATTT	4541	Db	1801	TCAGGGGCTAATAGGGAATAATACCATCTCTTTGCAAGCTTGGGTATCAAGTGAATTT	1860
Qy	4542	AACCACTAAGATTTATTTGTCGCTCAAGGCTGTTAAATTCAGAAAAAAGAAAGCGA	4601	Db	1861	AACCACTAAGATTTATTTGTCGCTCAAGGCTGTTAAATTCAGAAAAAAGAAAGCGA	1920
Qy	4602	AGCTCAACGTGTCATTTGTCAGAAATGAAAGAGATTTAATGCTTATATTACGGAATA	4661	Db	1921	AGCTCAACGTGTCATTTGTCAGAAATGAAAGAGATTTAATGCTTATATTACGGAATA	1980
Qy	4662	AGCGATGATACAGACCTTATTTAGTCAGCAACCAAAAGAGATTTAGAGAGTCTAGG	4721	Db	1981	AGCGATGATACAGACCTTATTTAGTCAGCAACCAAAAGAGATTTAGAGAGTCTAGG	2040
Qy	4722	CATTCTGCAACGCAATTAGATAAATTTGCTGAAGTACTGAAGCGGAATCAGGAAATTTT	4781	Db	2041	CATTCTGCAACGCAATTAGATAAATTTGCTGAAGTACTGAAGCGGAATCAGGAAATTTT	2100
Qy	4782	CTTTAGATTAACCAAGAGAAATGGTGGCATCTCACTCTAGTGTAAATCATTTGTT	4841	Db	2101	CTTTAGATTAACCAAGAGAAATGGTGGCATCTCACTCTAGTGTAAATCATTTGTT	2160
Qy	4842	GCTATCGATCAATTAAGTAAAAAAGAAAGAAAGAAAGCTATATAAGGCGCTGACAAA	4901	Db	2161	GCTATCGATCAATTAAGTAAAAAAGAAAGAAAGAAAGCTATATAAGGCGCTGACAAA	2220
Qy	4902	TTCTTTTGTAGCTAGAGCATACATTCATCAAGAGACTTTAAACAAGCTAGCAGAACGCC	4961	Db	2221	TTCTTTTGTAGCTAGAGCATACATTCATCAAGAGACTTTAAACAAGCTAGCAGAACGCC	2280
Qy	4962	TAAACGGACACACACTCGATTTGTTAGCTATGATACAGCGCTGAAATTAACACCGCA	5021	Db	2281	TAAACGGACACACACTCGATTTGTTAGCTATGATACAGCGCTGAAATTAACACCGCA	2340
Qy	5022	CTATGCCATTACATTTATATCATGATACGCTGTTTGTGTTTCTTCTGTTTATAGCGAAT	5081	Db	2341	CTATGCCATTACATTTATATCATGATACGCTGTTTGTGTTTCTTCTGTTTATAGCGAAT	2400
Qy	5082	GATTACGAGAAATATACAGATTAAGATTTTAAATTAATTTAGGGGAGAGAGAGAGT	5141	Db	2401	GATTACGAGAAATATACAGATTAAGATTTTAAATTAATTTAGGGGAGAGAGAGAGT	2460
Qy	5142	AGCCCGAAAACTTTTAGTTGGCTTGGACTGCAACGAAGTGAAGGAAAGGCTACTAAAAAGT	5201	Db	2461	AGCCCGAAAACTTTTAGTTGGCTTGGACTGCAACGAAGTGAAGGAAAGGCTACTAAAAAGT	2520
Qy	5202	CGAGGGGAGTGAAGCGAAGCAACCTGATTTTAAATTTCTATCTTTTATAGGTC	5261	Db	2521	CGAGGGGAGTGAAGCGAAGCAACCTGATTTTAAATTTCTATCTTTTATAGGTC	2580
Qy	5262	ATTAGAGTACTATTCTCTCTATAAATTTTAGCAGCATATAGATTTATTGAATAG	5321	Db	2581	ATTAGAGTACTATTCTCTCTATAAATTTTAGCAGCATATAGATTTATTGAATAG	2640
Qy	5322	GTCAATTAAGTTAGCATATTAGAGGAGGAAATCTTTGGAGAAATATTTCAGAAACCGCA	5381	Db	2641	GTCAATTAAGTTAGCATATTAGAGGAGGAAATCTTTGGAGAAATATTTCAGAAACCGCA	2700
Qy	5382	TTACATGGAATCGATTTCTTGTGGTTAGCTGTTTAACTAAAGTAGTGAATTTT	5441	Db	2701	TTACATGGAATCGATTTCTTGTGGTTAGCTGTTTAACTAAAGTAGTGAATTTT	2760
Qy	5442	TGATTTTGGTGTGTGCTGTTGTTGTAGTATTGCTAGTCAAGAGTAAATA	5497				

derived from GenBank Accession Number AF041239"		Db		7269 ACATTTATAAAATGTTTGACGACATTTTATAAATGCATAGCCGATAGATTGCCAAACC		7210	
/codon_start=1							
/transl_table=11							
/product="replication protein E"		QY		2964 AACGCTTACGTTAGTCAGATGAACTCTCCCTCGTAAGAAGTTATTTTAATTAACCTTTG		3023	
/protein_id="AAP73798.1"							
/db_xref="GI:32185268"		Db		7209 AACGCTTACGTTAGTCAGATGAACTCTCCCTCGTAAGAAGTTATTTTAATTAACCTTTG		7150	
/translation="MNIPFVVEVTLHDGLLKYKFKSKIRSIITKPKSKGKALFAYRS		QY		3024 TTTGAAGACGGGTATATAACCGTACTATCATTTATATAGGAAATCAGAGAGTTTCAAGTA		3083	
KSSMIGRGVLTSEBAIQNODTFHTWPNVRYGYADENSYTKHSENNLRQIN							
TFPDPDHPHAKETISASDILTPALDGFMEPTMLIKSDGQVAFVLETPVVTYSKSE		Db		7149 TTTGAAGACGGGTATATAACCGTACTATCATTTATATAGGAAATCAGAGAGTTTCAAGTA		7090	
FKSKAAKIQNIREIFGKSLPDLTCNHFGIARIPRTDNVEFFPNRYRYSFKWQD							
WSFQTKNKGFTSSLSVTEGKQVDSEFNNLLHETKFSGEKGLIGRNMMFTL		QY		3084 TCTAAGCTACTGAATTTAAGAAATTTGTTAAGCAATCAATCGGAAATCGTTTCATTCCTTTT		3143	
SLAYFSSGYETCEYNNFFNNRLDQLEKEVIKIVRSAYSENTOGANNREYITLLC							
KAWVSDDTSDLVRCQFKFKKSRQKRVHLSEWKEDLMAYISEKSDVYKPYLVT		Db		7089 TCTAAGCTACTGAATTTAAGAAATTTGTTAAGCAATCAATCGGAAATCGTTTCATTCCTTTT		7030	
TKREIRVLGTPERTSLKLVKANQEIPEFKIPGRNGGIQLASVKSILLSIIRVKK							
EEKESYIKALTNPSDLSEHTFIQETLNKLAERPKTDQLDLFSYDTG"		QY		3144 TTTGTATTTCATTTATAGAAGGTGGAGTTTGTATGAATCATGATGAATGAATCTATAT		3203	
/note="p1253 replication in gram positive bacteria;							
derived from GenBank Accession Number AF041239"		Db		7029 TTTGTATTTCATTTATAGAAGGTGGAGTTTGTATGAATCATGATGAATGTAAAACTTATAT		6970	
/codon_start=1							
/transl_table=11		QY		3204 AAAAAATAGTTTATGGAGATAAGAAAATTTAGCAAAATCTATACACTAGAGAAAGCTTTAA		3263	
/product="replication protein D"							
/protein_id="AAP73799.1"		Db		6969 AAAAAATAGTTTATGGAGATAAGAAAATTTAGCAAAATCTATACACTAGAGAGCTTTAA		6910	
/db_xref="GI:32185269"							
/translation="MNHDECKTYIKNSLLIRKLANIYTLTPKKLEKENIYLETKS		QY		3264 GAAAGATTGAGAAAGACAAATATCTACTTAGAAACAAATACAGTAAGTATTTTCTTC		3323	
DKYFSSGSEDIYKLIENNKIIYSIGKKLTYIKGKSFKAHLKQLNKAQNV"							
8339..8989		Db		6909 GAAAGATTGAGAAAGACAAATATCTACTTAGAAACAAATACAGTAAGTATTTTCTTC		6850	
/note="functions in gram positive bacteria; derived from		QY		3324 GGAGGGGGAAGATTATATATAAGTTTAATAGAAAATTAACAAAATAATTTATTCGATTAG		3383	
pC194 found in GenBank Accession Number V01277"							
/codon_start=1		Db		6849 GGAGGGGGAAGATTATATATAAGTTTAATAGAAAATTAACAAAATAATTTATTCGATTAG		6790	
/transl_table=11							
/product="chloramphenicol acetyltransferase"		QY		3384 TCGAAAAAATTTGACTTATAAGGAGAAAAATCTTTTCAAAACATGCAATATTGAAACA		3443	
/protein_id="AAP73800.1"							
/db_xref="GI:32185270"		Db		6789 TCGAAAAAATTTGACTTATAAGGAGAAAAATCTTTTCAAAACATGCAATATTGAAACA		6730	
/translation="MNFNKIDLDNMWRKEIFNHYLNQOFTTSITTEIDISLVYRNLIKQ		QY		3444 GTTGAATGAAAAAGCAACCAAGTTTAATTAACCAACCTATTTTATAGGATTATAGGAAA		3503	
EGYKFPAPFLVPLRVNSNTGYNDSGELGWDKLEPLTYIDGVSKEPSTGIWT							
PVKNDFEYDLVSDVEKYNKGSKLPKPTIPENAPLSLIIIPWTSTGFNLNNNS		Db		6729 GTTGAATGAAAAAGCAACCAAGTTTAATTAACCAACCTATTTTATAGGATTATAGGAAA		6670	
NYLLPITAGKFNKNSIYLPISLQVHVSVDGYHAGLFMNSIQBELSDRPNDWLL"							
complement(10617..11276)		QY		3504 GGAGAACGCTGAATGAATATCCCTTTTGTGTAGAACTGCTTCATGACGCGCTTGTT		3563	
/note="functions in gram negative bacteria; derived from							
plasmid pACYC184 found in GenBank Accession Number X06403"		Db		6669 GGAGAACGCTGAATGAATATCCCTTTTGTGTAGAACTGCTTCATGACGCGCTTGTT		6610	
/codon_start=1							
/transl_table=11		QY		3564 AAAGTACAAAATTTAAAAATAGTAAAAATTCGCTCAATCACTACCAAGCCAGGTAAAAAGCAA		3623	
/product="chloramphenicol acetyltransferase"							
/protein_id="AAP73801.1"		Db		6609 AAAGTACAAAATTTAAAAATAGTAAAAATTCGCTCAATCACTACCAAGCCAGGTAAAAAGCAA		6550	
/db_xref="GI:32185271"							
/translation="MEXKITGYTTVDISQWHRKEHFAFOSVACQTYNQTVQLDITAF		QY		3624 AGGGGCTATTTTGGCTATCGCTCAAAATCAAGCATGATTCGGCGGTCGCTGCTGTTGTTCT		3683	
LKYTKNKHFPYAPFIHILALNMNHPFPRMAMKQBELVLDWSVHPCTYVFHEQETFP							
SSLSEYHDDPQPLHYISQDVACYGENLAYFPKGFENNFFVSANPWSFTSFDLNV		Db		6549 AGGGGCTATTTTGGCTATCGCTCAAAATCAAGCATGATTCGGCGGTCGCTGCTGTTGTTCT		6490	
ANMDFAPVFTMGKIYTGQDKVLMPLAIQVHVAACDGFHVGRMLNLOQYCDWQGG							
A"		QY		3684 GACTTCCGAGGAGCGATTCAAGAAAATCAAGATACATTTTACATTTGGACACCCCAACGT		3743	
Query Match	51.1%; Score 2807.6; DB 12; Length 11925;						
Best Local Similarity	99.6%; Pred. No. 0;						
Matches 2824; Conservative	0; Mismatches						
	9; Indels						
	1; Gaps						
QY	2664 AGAACAAATTCGAAAGTTAACTCAACCTAGTAGTGATTATTAATCCCAATGAGCC	2723		3744 TTATCGTTATGGAAAGTATGAGACGAAAAACCGTTTCATACAGAAAGACATTCGAAAA	3803		
Db	7508 AGTATAAATTTAACGATCAACTCACTCAACCTAGTAGTGATTATTAATCCCAATGAGCC	7449		6429 TTATCGTTATGGAAAGTATGAGACGAAAAACCGTTTCATACAGAAAGACATTCGAAAA	6370		
QY	2724 AACAGAACGAGGCGAGAAACAGAAATCAGAACTGAATTAACATTTAGAAATGGAAGA	2783		3804 CAATTTAAGACAAAATCAATACCTTTTATTTGATTTCATATTCACACGGCAAAAGAAAC	3863		
Db	7448 AACAGAACGAGGCGAGAAACAGAAATCAGAACTGAATTAACATTTAGAAATGGAAGA	7389		6369 CAATTTAAGACAAAATCAATACCTTTTATTTGATTTCATATTCACACGGCAAAAGAAAC	6310		
QY	2784 AGAAAAAGCAATGACTTCGTGTGAATTAATGACGAAATTCGTTCTTATTTTTTTTAAA	2843		3864 TATTTTCAGCAACCGATATTTTAAACACCGCTATTGATTTCGCTTATTCGCTACTATGAT	3923		
Db	7388 AGAAAAAGCAATGACTTCGTGTGAATTAATGACGAAATTCGTTCTTATTTTTTTTAAA	7330		6309 TATTTTCAGCAACCGATATTTTAAACACCGCTATTGATTTCGCTTATTCGCTACTATGAT	6250		
QY	2844 AGCGGTATCTAGATATTAACGAAACCAACGAACTGAATAGAAACGAGAGAGCCATGAC	2903		3924 TATCAAACTGATAAGGTTTATCAAGCATATTTTGTTTTAGAAACGCCAGCTTATGTCAC	3983		
Db	7329 AGCGGTATCTAGATATTAACGAAACCAACGAACTGAATAGAAACGAGAGAGCCATGAC	7270		6249 TATCAAACTGATAAGGTTTATCAAGCATATTTTGTTTTAGAAACGCCAGCTTATGTCAC	6190		
QY	2904 ACATTTATAAATGTTTGACGACATTTTATAAATGCATAGCCCGATAGATTGCCAAACC	2963		3984 TTTCAAAATCAGAAATTTAAATCTGTCAAAAGCAGCCAAAATAATTTTCGAAAAATATCCGAGA	4043		

ORIGIN









ORGANISM Streptococcus pyogenes  
 Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 Streptococcus.  
 REFERENCE 1 (bases 1 to 9084)  
 AUTHORS Ceglowski, P., Boitsov, A., Chai, S. and Alonso, J.C.  
 TITLE Analysis of the stabilization system of pSM19035-derived plasmid  
 pB223 in *Bacillus subtilis*  
 JOURNAL Gene 136 (1-2), 1-12 (1993)  
 MEDLINE 94123985  
 PUBMED 8293991  
 REFERENCE 2 (bases 1 to 9084)  
 AUTHORS Alonso, J.C.  
 TITLE Direct Submission  
 JOURNAL Submitted (27-FEB-1992) J.C. Alonso, Max-Planck-Institut fuer  
 Molekulare Genetik, Innestr.73, D-1000 Berlin 33, FRG  
 FEATURES  
 source  
 1..9084  
 /organism="Streptococcus pyogenes"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:1314"  
 /plasmid="pB223"  
 135..163  
 /note="Pli promoter"  
 503..1993  
 /gene="reps"  
 503..1993  
 /gene="reps"  
 /codon\_start=1  
 /transl\_table=1  
 /protein\_id="CAA45928.1"  
 /db\_xref="GI:456363"  
 /db\_xref="GOA:P14752"  
 /db\_xref="SWISS-PROT:P14752"  
 /translation="MNIPFVETVHGLKLYFKSKIRSIITTKPKSGKALFAYRS  
 KXSLGRGVLTSEEAHNEQDFTHTVNVRYGYADENSYKSGSENNLRQIN  
 TFFIDFHEKTSADILITTAIDLGFMPTLIIISDKGYQFYVLETPVYVTSKSE  
 FSKVQAQKISQNIRESYFGKSLPVDLTCNHFGLIARIPRTDNVEFFDPNRYSPKRWQD  
 WSFQTKQKISQNIRESYFGKSLPVDLTCNHFGLIARIPRTDNVEFFDPNRYSPKRWQD  
 LAWFSYSGYSETCEYNNFNNRLDQLEKEVILKVRSAYSENTQGANREYITLIC  
 KAVVSDLTSDKDFVQGFKKRSEKORVHLSEWKEDLMAYISEKSDVYKPLVT  
 TKELREALGIPETTLKLLKLVKANOEIFPKIKSGRNGGIGLASGKSLILLSIKVKK  
 EKESYIKALTNFSDELHTFIQETLNLKLRPKTDQDLDFSYDTG"  
 2341..2511  
 /note="unnamed protein product; ORF alpha"  
 /codon\_start=1  
 /transl\_table=1  
 /protein\_id="CAA45929.1"  
 /db\_xref="GI:456364"  
 /db\_xref="GOA:Q57042"  
 /db\_xref="SWISS-PROT:Q57042"  
 /translation="MAKIGYARVSKQNLDRLOALQGVSKVFSKLSQGVSRPOL  
 QAMLYNREGDIVVTVLDRGRNKKELTLMNAIQKQATLEVLNLPWMNGEDENL  
 RLLNNVILEYKYQAESERKRIKEROAQIEIAKSKGPKGQHKFENDPLKHP  
 DFLNGSDKVEQETGINTRTFRYTRVTVQKNGKSDS"  
 3142..5286  
 /note="unnamed protein product; ORF gamma"  
 /codon\_start=1  
 /transl\_table=1  
 /protein\_id="CAA45931.1"  
 /db\_xref="GI:456365"  
 /db\_xref="GOA:Q60149"  
 /db\_xref="SWISS-PROT:Q60149"  
 /translation="MSTVILAEKPSQALAYAQAFNQSDKDKGYFEIKDPLFTDTEFIT

EFGLHVELAEAGNYDEKQWKNLESLEIPDRDYDFEVAKDGKQFKI VAEALLKKANT  
 LIVATDSREGENIAMSIIHKANAFSKDKTKFKRLWINSLEKDVIRSGFQNLQPCMNYY  
 PFYQEQTRQIADLIGNASPLTYLNLNQKGVQGTSLGRVQTPTYLILFQREAEIE  
 NFKPEFFPEVEASIKVNGSPKVLSPQRFKTEQELLAFVSSKQAKIGQEGKITDV  
 QTKKKTNSPSLFSLSQKVNQLYKATASQTLKAMQGVYEAKLLSPRTDTPFITE  
 NEFYAKANFGKYGSLGLDLEMTQPRKYVDGSKVQEHAIIPKQVPTESALAK  
 MDDLQKTIYLVVKTIVAMPDLYLVEETKIOTKADVLLFOSIGKTPKQKWLIFKQ  
 QNKKEDEVOTPLVILGEAEVDVSAKETOPPKAFTEGTLTAKTAKGTIDDBE  
 AQKILKQTEGTEATRASIIEALKQXEYIQVILKLVKTEKGLLQCAVESOHLLES  
 AEMTAKMETYLKIKRGEGNQENFITNIKFIHLLAEAVPTDIEKLNFSYQEQKEXE  
 AEKIVGKCPKCGNNIVLKSYGCSNYPECKFTLAEHFRKKKLTNTNVKLELGEKET  
 LVKGIKKEKSKSYNAVYKIGKYDIFISFK"  
 5389..6285  
 /note="unnamed protein product; ORF delta"  
 /codon\_start=1  
 /transl\_table=1  
 /protein\_id="CAA45932.1"  
 /db\_xref="GI:456367"  
 /db\_xref="GOA:Q57280"  
 /db\_xref="SWISS-PROT:Q57280"  
 /translation="MIQYVYTKGVMKEELKLEELRILSNKNEAIVILNNYF  
 KGVGKSKLSTWFAVLDKLNKVMIDKDIQATLTAKTFKVLPRVNFYEGLNK  
 GNLAASSIVHLTDNLDPGTPDMLLPKLTFSWTENESRLATLLAPLKSDDYLLII  
 DTVTPPSVTNNAIVASDVIWIPQAEESTNNIYISYILIDQEQFNGCLDNLGV  
 PVLVDTSATIKSLEELYKQKEDNLVFNQIIRSNKSVTWSKNGKITEHKGDKKVL  
 SMYKNVFFEMLERIQLENEKE"  
 6609..6881  
 /note="unnamed protein product; ORF epsilon"  
 /codon\_start=1  
 /transl\_table=1  
 /protein\_id="CAA45933.1"  
 /db\_xref="GI:456368"  
 /db\_xref="GOA:Q57231"  
 /db\_xref="SWISS-PROT:Q57231"  
 /translation="MAVTEKTFEIEIINELASVYNNVNLNHLNKNDSQLEEV  
 NLLNQKLAKENVLFDYSLEQLQAVEHWSMNSYKQVLNKEVA"  
 6883..7746  
 /note="unnamed protein product; ORF zeta"  
 /codon\_start=1  
 /transl\_table=1  
 /protein\_id="CAA45934.1"  
 /db\_xref="GI:456369"  
 /db\_xref="GOA:Q54944"  
 /db\_xref="SWISS-PROT:Q54944"  
 /translation="MANIVFTDKQFNRLNDNLEELIQKKAVERPTAELLGQPGS  
 KTSLSAIFETQGNVINDITFKQHPNFDLKLKYEDKVVHVTPSYNRMTEAI  
 ISRLSDQGNLVIETGRTDVPITQATMLQAGYETKVMYVAFKINSYGLTERYE  
 TWYADDPMTARATPKQADHIVVKNLPTNLETHTGLFSIRLYNREGVKLSLETP  
 SISPKETLEKLNKVSCKEIQPTLERIEQKWLNKIQTPEFKAIOQLKLESLOPTP  
 PIPKTPKLPGI"  
 complement (8009..8746)  
 /gene="erm2"  
 complement (8009..8746)  
 /gene="erm2"  
 /codon\_start=1  
 /transl\_table=1  
 /protein\_id="CAA45935.1"  
 /db\_xref="GI:456370"  
 /db\_xref="GOA:Q54945"  
 /db\_xref="SWISS-PROT:Q54945"  
 /translation="MNKNIKYSQNFLEKVLNQIILKQNLKETTDTVYETGTGKHLT  
 TKLAKISQVTSIELDSHFLNLSKLNIRVLIHQDILQFQPNKQYKIVGNIP  
 YHLSTQIILKVVVFESHASDIYIVVEGFYKRTDLHRTGLLHTQVSIQOLLKPAE  
 CFHPKQVNSVLIKLTHTTDDVPKYNKLYTFYFVSKWVNRVYRQLFTKNQHPQAKHA  
 KYNLSVTYEQVLSIFNSYLLFNGRK"  
 complement (8839..8949)  
 /gene="erm1"  
 complement (8839..8949)  
 /gene="erm1"  
 /codon\_start=1  
 /transl\_table=1  
 /protein\_id="CAA45936.1"  
 /db\_xref="GI:456371"

/translation="MLVFQMRNVNDKTSVLEKQTKNSDYADKVRLLPTSD"									
ORIGIN									
Query Match 31.6%; Score 1735.4; DB 1; Length 9084;									
Best Local Similarity 95.4%; Pred. No. 3.8e-270;									
Matches 1798; Conservative 0; Mismatches 86; Indels 1; Gaps 1;									
QY	3488	TAGATTTATAGAAAGGAGAAACAGCTGAATGATATCCCTTTTGTGTAGAAACGTGCG	3547						
DB	474	TAATTTTCGTAAGAAAGGAGAAACAGCTGAATGATATCCCTTTTGTGTAGAAACGTGCG	533						
QY	3548	TTCAATGACGGCTTTGTTAAAGTACAAATTTAAATAATAGTAAAAATTCGCTCAATCACTACCA	3607						
DB	534	TTCAATGACGGCTTTGTTAAAGTACAAATTTAAATAATAGTAAAAATTCGCTCAATCACTACCA	593						
QY	3608	AGCAGGTAAAGCAAGGGGCTATTTTTCGGTATCGCTCAAAATCAAGCATGATGGCG	3667						
DB	594	AGCAGGTAAAGCAAGGGGCTATTTTTCGGTATCGCTCAAAATCAAGCATGATGGCG	653						
QY	3668	GTGCTGTGTGTTCTGTACCTTCGAGGAGCGGATTCAGAGAAATCAAGATACATTTACAC	3727						
DB	654	GAGTGGCGTGTGTTCTTACCTTCGAGGAGCGGATTCAGAGAAATCAAGATACATTTACCG	713						
QY	3728	ATTGGACACCAACGTTTATCGTTATGGAAACGATGACAGACGAAACCGGTTTACACGA	3787						
DB	714	ATTGGACACCAACGTTTATCGTTATGGTACGTTATGACAGGAAACCGGTTTACACCTA	773						
QY	3788	AGGACATTCCTGAACCAATTTAAGCAAAATCAATACCTTCTTTATGATTTGATATTC	3847						
DB	774	AGGACATTCCTGAACCAATTTAAGCAAAATCAATACCTTCTTTATGATTTGATATTC	833						
QY	3848	ACACGGCAAAAGAACTATTTTACGACGCGATTTTAAACACCGCTATTTGATTTAGGTT	3907						
DB	834	ACACGGCAAAAGAACTATTTTACGACGCGATTTTAAACACGCTATTTGATTTAGGTT	893						
QY	3908	TTATGCTACTATGATTTCAAAATCTGATAAGGTTATCAAGCATATTTTGTGTTAGAA	3967						
DB	894	TTATGCTACTATGATTTCAAAATCTGATAAGGTTATCAAGCATATTTTGTGTTAGAA	953						
QY	3968	CGCCAGTCTATGTGACTTTCAAAATCAGATTTAAATCTGCAAGCAGCCCAAAATATTT	4027						
DB	954	CGCCAGTCTATGTGACTTTCAAAATCAGATTTAAATCTGCAAGCAGCCCAAAATATCT	1013						
QY	4028	CGCAAAATATCCGAGAAATTTTGGAAAGCTTTTGGCAGTTGATCTAAACGTTGATCAAT	4087						
DB	1014	CGCAAAATATCCGTTGAATTTTGGAAAGCTTTTGGCAGTTGATCTAAACGTTGATCAAT	1073						
QY	4088	TTGCTATTTCTGCTATACCAAGACGACAAATGTTAGAAATTTTGTGCTTCTTAATTCGGTT	4147						
DB	1074	TTGGGATTTGCTGCTATACCAAGACGACAAATGTTGAAATTTTGTGCTTCTTAATTCGGTT	1133						
QY	4148	ATTCTTTCAAGAAATGCGAAGATTGGTCTTTTCAAAACAAACAGATAATAAGGGCTTTACTC	4207						
DB	1134	ATTCTTTCAAGAAATGCGAAGATTGGTCTTTTCAAAACAAACAGATAATAAGGGCTTTACTC	1193						
QY	4208	GTTCAAGTCTAAACGTTTAAAGGGTACAGAGGCAAAACAAAGTAGATGAACCTGGT	4267						
DB	1194	GTTCAAGTCTAAAGGTTTAAAGGGTACAGAGGCAAAACAAAGTAGATGAACCTGGT	1253						
QY	4268	TTAATCTCTTATTCGACGAAACGAAATTTTCAAGAGAAAGGGTTTAAATAGGGCGTAATA	4327						
DB	1254	TTAATCTCTTATTCGACGAAACGAAATTTTCAAGAGAAAGGGTTTAAATAGGGCGTAATA	1313						
QY	4328	ACGTCAATGTTTACCTCTCTTTAGCCTACTTTAGTTCAGGCTATTCATCAAGACGTGCG	4387						
DB	1314	CGGTATGTTTACCTCTCTTTAGCCTACTTTAGTTCAGGCTATTCATCAAGACGTGCG	1373						
QY	4388	AATATAATGTTTGGTGTAAATATCGATTAGATCAACCTTTAGAGAAAGAAAGTAA	4447						
DB	1374	AATATAATGTTTGGTGTAAATATCGATTAGATCAACCTTTAGAGAAAGAAAGTGA	1433						
QY	4448	TCAAAATTTGTTAGAGTGCCTTATTCAGAAACCTATCAAGGGGCTTAATAGGGAATACATTA	4507						

DB	1434	TCAAACTTGTGTAGAGTGCCTACTCAGAAACTATCAAGGGCTAATAGGGAATACATTA	1493
QY	4508	CCATCTTTTCAAAGCTTGGTATCAAGTGATTTAAACAGTAAAGATTTATTTGTCCGTC	4567
DB	1494	CCATCTTTTCAAAGCTTGGTATCAAGTGATTTAAACAGTAAAGATTTATTTGTCCGTC	1553
QY	4568	AAGGTGTGTTTAAATTTCAAGAAAAAGAGCAAGCTCAACGTTTCAATTTGTGAGAAAT	4627
DB	1554	AAGGTGTGTTTAAATTTCAAGAAAAAGAGCAAGCTCAACGTTTCAATTTGTGAGAAAT	1613
QY	4628	GGAAAGAGATTTTAAATGCTTTATTTAGCGAAAAAGAGATGTATACAAAGCTTTATTAG	4687
DB	1614	GGAAAGAGATTTTAAATGCTTTATTTAGCGAAAAAGAGATGTATACAAAGCTTTATTAG	1673
QY	4688	TCAGACCAAAAGAGATTTAGAGAGTGTAGGCAATTCCTGAAACGCAATTTAGATAAT	4747
DB	1674	TCAGACCAAAAGAGATTTAGAGAGGCTTAGGCAATTCCTGAAACGCTAGATAAGC	1733
QY	4748	TGCTGAAGTACTGAGGCGGAATCAGGAAATTTTCTTTAAAGATTTAAACACGAGAAATG	4807
DB	1734	TATTGAAGTATTAAACGGAATCAAGAAATCTTCTTTAAGATTTAAATCAGAGAAATG	1793
QY	4808	GTGGCATTCATCTGCTAGTGTAAATCATTTGTTGCTATCGATCATTTAAAGTAAAGAAAG	4867
DB	1794	GTGGCATTCATCTGCTAGTGTAAATCATTTGTTGCTATCGATCATTTAAAGTAAAGAAAG	1853
QY	4868	AAGAAAAAGAGCTATATAAAGGCGCTGACAAATTTCTTTGACTTAGAGCATACATTCA	4927
DB	1854	AAGAAAAAGAGCTATATAAAGGCGCTGACAAATTTCTTTGACTTAGAGCATACATTCA	1913
QY	4928	TTCAAGAGACTTTAAACAGCTAGCAGAACCCCTTAAACCGGACACACAACTCGATTTGT	4987
DB	1914	TTCAAGAGACTTTAAACAGCTAGCAGAACCCCTTAAACCGGACACACAACTCGATTTGT	1973
QY	4988	TTAGCTATGATACAGGCTGAAAAATAAACCCGCACTATGCCATTACATTTATCTATGA	5047
DB	1974	TTAGCTATGATACAGGCTGAAAAATAAACCCGCACTATGCCATTACATTTATCTATGA	2033
QY	5048	TACGTGTTTGTGTTTCTTTGCTGTTAGCGAATGATTAGCAGAAATATACAGATAAGA	5107
DB	2034	TACGTGTTTGTGTTTCTGCTGTTAGTGAATGATTAGCAGAAATATAGAGATAAGA	2092
QY	5108	TTTTTAATTAATTTAGGGGAGAGAGAGTAGCCCGAAACTTTTGTGCTTTGG	5167
DB	2093	TTTTTAATTAATTTAGGGGAGAGAGAGTAGCCCGAAACTTTTGTGCTTTGG	2152
QY	5168	ACTGAAACGAAGTGAAGGAAAGGCTACTAAACGTCGAGGGGCGAGTGAAGCGGAAC	5227
DB	2153	ACTGAAACGAAGTGAAGGAAAGGCTACTAAACGTCGAGGGGCGAGTGAAGCGGAAC	2212
QY	5228	ACTGATTTTAAATTTCTATCTTTTATAGGTCAATTAGAGTATCTTTTGTGCTATA	5287
DB	2213	ACTGATCTTTTAAAGTTGCTATCTTTTATAGGTCAATTAGGTATCTTTTGTGCTATT	2272
QY	5288	AACTATTTTAGCAGCATATAATAGATTATTGTAATAGGTCAATTAAAGTTGAGCATATTAGAGG	5347
DB	2273	GAATTAGATAGCATATAATAGCTTTTATAGGTAGGTCAATTAAAGTTGAGCATATTAGGAG	2332
QY	5348	AGGAAATCTTTGGAGAAATATTGA	5372
DB	2333	GATCAAGAAATGAAAAATTTTATTA	2357

Search completed: September 1, 2004, 20:41:02  
Job time : 20461 secs

This Page Blank (uspio)

OM nucleic - nucleic search, using sw model

Run on: September 1, 2004, 13:08:39 ; Search time 1889 Seconds  
(without alignments)  
12362.279 Million cell updates/sec

Title: US-10-030-390-2  
Perfect score: 5497  
Sequence: 1 gaattgattagtcattctt.....ctagtcacaaagtattataata 5497

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:.\*  
1: Geneseqn1980s:.\*  
2: Geneseqn1990s:.\*  
3: Geneseqn2000s:.\*  
4: Geneseqn2001as:.\*  
5: Geneseqn2001bs:.\*  
6: Geneseqn2002s:.\*  
7: Geneseqn2003as:.\*  
8: Geneseqn2003bs:.\*  
9: Geneseqn2003cs:.\*  
10: Geneseqn2004s:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5497	100.0	5497	5	Aaf29646 Plasmid p
2	5010.8	91.2	5770	3	AAA15043 Nucleotid
3	5001.4	91.0	5906	3	AAA15042 Nucleotid
4	4999	90.9	5870	3	AAA15044 Nucleotid
5	4993.4	90.8	5230	3	Aax07549 pTREX1 mu
6	4991.8	90.8	5230	3	AAA15041 Nucleotid
7	4982.4	90.6	5216	2	Aax07548 pK930 IIR
8	4944.6	90.0	5231	2	Aax07550 pTREXIA m
9	1434.8	26.1	3765	2	AAA13224 Enterococ
10	1434.8	26.1	3765	6	ABS99019 Enterococ
11	1111	20.2	10323	2	Aax09054 Plant tra
12	1111	20.2	17111	3	Aad01289 Arabidops
13	1111	20.2	17116	3	Aad01290 Arabidops
14	1109.4	20.2	10929	6	ABK98591 PER25 vec
15	1109.4	20.2	10929	8	ACD13842 Plasmid p
16	983.6	17.9	15676	3	Aad01288 Arabidops
17	884	16.1	6243	2	Aaz09598 Clostridi
18	876.6	15.9	9412	6	ABK98635 Vector pE
19	876.6	15.9	9412	6	ABK98636 Vector pE
20	876.6	15.9	9412	8	ACD13886 L. lactis
21	876.6	15.9	9412	8	ACD13887 L. lactis
22	876.6	15.9	9417	6	ABK98632 Vector pE
23	876.6	15.9	9417	8	ACD13883 L. lactis

C 24	876.6	15.9	9425	6	ABK98630	Vector pE
C 25	876.6	15.9	9425	8	ACD13881	L. lactis
C 26	876.6	15.9	9437	6	ABK98601	Vector pE
C 27	876.6	15.9	9437	8	ACD13852	Xyl opera
C 28	876.6	15.9	9484	6	ABK98633	Vector pE
C 29	876.6	15.9	9484	8	ACD13884	L. lactis
C 30	876.6	15.9	9566	6	ABK98634	Vector pE
C 31	876.6	15.9	9566	8	ACD13885	L. lactis
C 32	876.6	15.9	12733	6	ABK98631	Vector pE
C 33	876.6	15.9	12733	8	ACD13882	L. lactis
C 34	876.6	15.9	12739	6	ABK98592	Vector pE
C 35	876.6	15.9	12739	8	ACD13843	Plasmid p
C 36	833.8	15.2	9749	2	AAI13091	Enterococ
C 37	833.8	15.2	9749	6	ABS98886	Enterococ
C 38	747	13.6	747	6	ABK78393	Bacillus
C 39	731.6	13.3	738	2	AAV37113	Antibioti
C 40	731.6	13.3	738	4	AAH01718	Escherich
C 41	471	8.6	5142	5	AAF29645	Plasmid p
C 42	381	6.9	381	6	ABK73933	Bacillus
C 43	295.6	5.4	4995	7	AAI51077	pothy16 p
C 44	295.6	5.4	6964	7	AAI51073	Lactococc
C 45	293	5.3	6904	7	AAI51072	Lactococc

ALIGNMENTS

RESULT 1  
AAF29646  
ID AAF29646 standard; DNA; 5497 BP.  
XX

AC AAF29646;  
XX

DT 10-APR-2001 (first entry)  
XX

DE Plasmid pTMTFF1.  
XX

XX Mouse TFF1; mTFF; trefoil protein; antiinflammatory; gastrointestinal;  
XX anti-ulcer; peptide therapy; gastrointestinal disease; acute colitis;  
KW Crohn's disease; ulcerative colitis; plasmid pTMTFF1;  
KW recombinant vector; ds.  
XX

OS Synthetic.  
XX

XX WC200102570-A1.  
FN

XX 11-JAN-2001.  
PD

XX 05-JUL-2000; 2000WO-EP006343.  
PF

XX 05-JUL-1999; 99EP-00870143.  
PR

XX (VLA-- ) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.  
PA

XX Hans WC, Steidler L, Remaut ER;  
PI

XX WPI; 2001-138142/14.  
DR

XX Recombinant Lactococcus lactis for delivering a trefoil peptide useful  
XX for treating acute or chronic gastrointestinal inflammatory diseases or  
PT disorders, e.g. acute or ulcerative colitis, acute flare-ups of Crohn's  
PT disease.  
XX

XX Claim 18; Fig 1c; 44pp; English.  
PS

XX The present sequence is a recombinant vector which may be used to  
CC generate a recombinant Lactococcus lactis capable of delivering a trefoil  
CC peptide in vivo. The recombinant microorganism is useful for  
CC manufacturing an agent for the delivery of a trefoil peptide to the  
CC gastrointestinal tract, and for treating gastric or intestinal diseases  
CC or disorders, or lesions caused by gastric or intestinal diseases or  
CC disorders. The microorganism may also be used for preparing medicament to  
CC be used for treating gastric and /or gastrointestinal diseases or

CC disorders, acute gastrointestinal inflammatory diseases (e.g., acute  
 CC colitis, acute flare-ups of Crohn's diseases, or ulcerative colitis), and  
 CC chronic and spontaneously recurring diseases of the gastrointestinal  
 CC tract comprising Crohn's disease (enteritis regionalis) and ulcerative  
 CC colitis (colitis ulcerosa). Disease states which can be treated by the  
 CC method or compositions comprising the recombinant microorganism or  
 CC trefoil peptides include disorders of and damage to the alimentary canal,  
 CC including the mouth, esophagus, stomach and large and small intestine,  
 CC as well as for the protection and treatment of tissues that lie outside  
 CC the alimentary canal  
 XX  
 SQ Sequence 5497 BP; 1994 A; 870 C; 1051 G; 1582 T; 0 U; 0 Other;

Query Match 100.0%; Score 5497; DB 5; Length 5497;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 5497; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GAATTCGATTAAGTCACTTACCTCTTTTATTAGTGTCTTTCTATAATCTAATGATAACA	60
DB	1	GAATTCGATTAAGTCACTTACCTCTTTTATTAGTGTCTTTCTATAATCTAATGATAACA	60
QY	61	TTTTTATAATTAATCTATTAACCATATCCCTCTTTGGAAATGAAATTTATTATCTACTCC	120
DB	61	TTTTTATAATTAATCTATTAACCATATCCCTCTTTGGAAATGAAATTTATTATCTACTCC	120
QY	121	TTTGTAGATATCTTATAATCAAGTATCAGATCTGGGAGACCAACCGTTTCCACTAG	180
DB	121	TTTGTAGATATCTTATAATCAAGTATCAGATCTGGGAGACCAACCGTTTCCACTAG	180
QY	181	AAATAATTTTGTAACTTTAGAAAGGAGATATACGATGAAAGAAAGATTTATCTCAGC	240
DB	181	AAATAATTTTGTAACTTTAGAAAGGAGATATACGATGAAAGAAAGATTTATCTCAGC	240
QY	241	TATTTTAAATGCTACAGTCAATCTTCTGCTGCGACCCCGTTGTCAGGTGTTTACGCCCA	300
DB	241	TATTTTAAATGCTACAGTCAATCTTCTGCTGCGACCCCGTTGTCAGGTGTTTACGCCCA	300
QY	301	GGCCAGGCCAGGCCAGGCCAGGAAACATATGATATGCGCCCGCCCGGGAGAGAT	360
DB	301	GGCCAGGCCAGGCCAGGCCAGGAAACATATGATATGCGCCCGCCCGGGAGAGAT	360
QY	361	AAATTTGCTGCTCCCGGTGTCACCGCCAGAGTGCAGCGAGAGAGGTGCTGTTTGA	420
DB	361	AAATTTGCTGCTCCCGGTGTCACCGCCAGAGTGCAGCGAGAGAGGTGCTGTTTGA	420
QY	421	TGACAGTGTCCGGGATTTCCGCTGTGTCTCCACCCATGCGCCATCGAGAACACTCAAGA	480
DB	421	TGACAGTGTCCGGGATTTCCGCTGTGTCTCCACCCATGCGCCATCGAGAACACTCAAGA	480
QY	481	AGAAGAAATGCTCTTAATCTAGTAGATCCGGCTGCTAAAGAGAGAGAGCTGA	540
DB	481	AGAAGAAATGCTCTTAATCTAGTAGATCCGGCTGCTAAAGAGAGAGAGCTGA	540
QY	541	GTGCTGCTGCTCCAGCGCTGAGCAATTAAGTAAACCCCTTGGGCGCTCTTAAACCGGT	600
DB	541	GTGCTGCTGCTCCAGCGCTGAGCAATTAAGTAAACCCCTTGGGCGCTCTTAAACCGGT	600
QY	601	CTTGAGGGGTTTTTGTCTGAAGAGAGAACTATATCCGGATGACCTCGAGCAAGCTCTA	660
DB	601	CTTGAGGGGTTTTTGTCTGAAGAGAGAACTATATCCGGATGACCTCGAGCAAGCTCTA	660
QY	661	GAAATCGATACGATTTTGAAGTGGCAACAGATATAAAGAGAGTTTAAATTTGTGCTGA	720
DB	661	GAAATCGATACGATTTTGAAGTGGCAACAGATATAAAGAGAGTTTAAATTTGTGCTGA	720
QY	721	ACTTTTAAAAAACAAGCAATCAATCTATGTCGCAACAGATAGCGACAGAGAGCGGAAAA	780
DB	721	ACTTTTAAAAAACAAGCAATCAATCTATGTCGCAACAGATAGCGACAGAGAGCGGAAAA	780
QY	781	CATTGCTGCTGCTGATCTTCAATGAAGCAATGCTTTTCTTAAAGATATAAAGCTATAAAG	840
DB	781	CATTGCTGCTGCTGATCTTCAATGAAGCAATGCTTTTCTTAAAGATATAAAGCTATAAAG	840

QY	841	ACTATGATCAATAGTATTAGAAAAAGATGTATCCGTAGCGGTTTTCAAAATTTGCAACC	900
DB	841	ACTATGATCAATAGTATTAGAAAAAGATGTATCCGTAGCGGTTTTCAAAATTTGCAACC	900
QY	901	AGGAATGAATTAATCTATCCCTTTTATCAAGAGCGGCAAAAGAAAAAGAAATGATACCA	960
DB	901	AGGAATGAATTAATCTATCCCTTTTATCAAGAGCGGCAAAAGAAAAAGAAATGATACCA	960
QY	961	ATCAGTGCAAAAAAGATATAATGGAGATAAGACGGTTTCGTTGCTGCTGCTGACATGCAC	1020
DB	961	ATCAGTGCAAAAAAGATATAATGGAGATAAGACGGTTTCGTTGCTGCTGCTGACATGCAC	1020
QY	1021	CATATCATTAATTAATCGAAACGCAAGATGCGGAAAGCTAAAGAAAGTATGGAATA	1080
DB	1021	CATATCATTAATTAATCGAAACGCAAGATGCGGAAAGCTAAAGAAAGTATGGAATA	1080
QY	1081	AGACTTAGAAGCAAACTTAAGAGTGTGTTGATAGTGCAGTATCTTAAATTTTGTATAAT	1140
DB	1081	AGACTTAGAAGCAAACTTAAGAGTGTGTTGATAGTGCAGTATCTTAAATTTTGTATAAT	1140
QY	1141	AGGAATTTGAAGTAAATTAAGATGCTTAAATTTTGTAAATTTGAAGAGAGTATGATGAA	1200
DB	1141	AGGAATTTGAAGTAAATTAAGATGCTTAAATTTTGTAAATTTGAAGAGAGTATGATGAA	1200
QY	1201	CAAAAAATATAAATATTCTCAAAACCTTTTAAACGAGTGAAGAAAGTACTCAACCAATTAAT	1260
DB	1201	CAAAAAATATAAATATTCTCAAAACCTTTTAAACGAGTGAAGAAAGTACTCAACCAATTAAT	1260
QY	1261	AAAAATTTGAATTTAAAGAAACCGATACCGTTTACGAAATTTGAAACAGGTAAAGGCA	1320
DB	1261	AAAAATTTGAATTTAAAGAAACCGATACCGTTTACGAAATTTGAAACAGGTAAAGGCA	1320
QY	1321	TTTAAAGCAAGAACTGGCTAAATTAAGTAAACAGGTAAAGTCTTAAATTTGAGACAGTCA	1380
DB	1321	TTTAAAGCAAGAACTGGCTAAATTAAGTAAACAGGTAAAGTCTTAAATTTGAGACAGTCA	1380
QY	1381	TCTATTCAACTTATCGTCAGAAAAATTTAAACTGAAATCTCGTGTCACTTTAAATTCACCA	1440
DB	1381	TCTATTCAACTTATCGTCAGAAAAATTTAAACTGAAATCTCGTGTCACTTTAAATTCACCA	1440
QY	1441	AGATATTTACAGTTTCAATTTCCCTAAACAGAGAGTATAAATTTGCGAGTATTTCC	1500
DB	1441	AGATATTTACAGTTTCAATTTCCCTAAACAGAGAGTATAAATTTGCGAGTATTTCC	1500
QY	1501	TTACCAATTTAAGCACACAAATTTATTTAAAAAGTGGTTTTTGAAGCCATCGCTCTGACAT	1560
DB	1501	TTACCAATTTAAGCACACAAATTTATTTAAAAAGTGGTTTTTGAAGCCATCGCTCTGACAT	1560
QY	1561	CTATCTGATTTTGAAGAGGATTTCTACAGGTACCTTGGATTTTACCCGCACTAGG	1620
DB	1561	CTATCTGATTTTGAAGAGGATTTCTACAGGTACCTTGGATTTTACCCGCACTAGG	1620
QY	1621	GTGCTCTTGACACCTCAAGTCTCGAATTTGCAATTTGCTTAAGCTGCGAGCGGATGCTT	1680
DB	1621	GTGCTCTTGACACCTCAAGTCTCGAATTTGCAATTTGCTTAAGCTGCGAGCGGATGCTT	1680
QY	1681	TGATCTTAAACCAAGTAAACAGTGTCTTAAATAAACTTTACCCGCCATACACAGATGT	1740
DB	1681	TGATCTTAAACCAAGTAAACAGTGTCTTAAATAAACTTTACCCGCCATACACAGATGT	1740
QY	1741	TCAGATAAATTTGGAAGCTATATACGTACTTTTCAAAATGGGTCAATTCAGAGATA	1800
DB	1741	TCAGATAAATTTGGAAGCTATATACGTACTTTTCAAAATGGGTCAATTCAGAGATA	1800
QY	1801	TCGTCAACTGTTTACTTAAAAATCAAGTTTTCATCAAGCAATGAAACACGCCAAGTAAACAA	1860
DB	1801	TCGTCAACTGTTTACTTAAAAATCAAGTTTTCATCAAGCAATGAAACACGCCAAGTAAACAA	1860
QY	1861	TTTAAAGTACCGTTACTTATGAGAGTATTCGTCTATTTTAAATAGTTATCTATTATTAA	1920
DB	1861	TTTAAAGTACCGTTACTTATGAGAGTATTCGTCTATTTTAAATAGTTATCTATTATTAA	1920
QY	1921	CGGAGGAGAAATTAATTTCTATGATCGCTTTTGTAAATTTGGAAGTATCACGTTACTAAAG	1980







Query Match				91.2%; Score 5010.8; DB 3; Length 5770;			
Best Local Similarity				100.0%; Pred. No. 0;			
Matches 5002; Conservative				0; Mismatches 0; Indels 0; Gaps 0;			
QY	496	CTAACTAGTAGATCCGGCTGCTAACAAAGCCGGAAGAGCACTGAGTTGGCTGCTGCAC	555				
DB	769	CTAACTAGTAGATCCGGCTGCTAACAAAGCCGGAAGAGCACTGAGTTGGCTGCTGCAC	828				
QY	556	CGCTGAGCAATACTAGCATAAACCCCTTTGGGGCCCTTAAACGGGTCTTGAGGGGTTTTT	615				
DB	829	CGCTGAGCAATACTAGCATAAACCCCTTTGGGGCCCTTAAACGGGTCTTGAGGGGTTTTT	888				
QY	616	GCTGAAGAGGAGCACTATATCCGGATGACCTGCGAGCAAGCTCTAGAAATCGATACGATTT	675				
DB	889	GCTGAAGAGGAGCACTATATCCGGATGACCTGCGAGCAAGCTCTAGAAATCGATACGATTT	948				
QY	676	TGAAGTGGCAACAGATAAAAAAGCAGTTTAAAAATTTGTTGCTGAACTTTTAAAAACAAGC	735				
DB	949	TGAAGTGGCAACAGATAAAAAAGCAGTTTAAAAATTTGTTGCTGAACTTTTAAAAACAAGC	1008				
QY	736	AAATACAAATCATTTGTCGAACAGATACCGACAGAGAGGCGAAACATTTGCTGGTGCAT	795				
DB	1009	AAATACAAATCATTTGTCGAACAGATACCGACAGAGAGGCGAAACATTTGCTGGTGCAT	1068				
QY	796	CATTCTAAAGCAAAATGCCCTTTTCTAAAGATAAAAACGTTATATAAAGACTATGGATCAATAG	855				
DB	1069	CATTCTAAAGCAAAATGCCCTTTTCTAAAGATAAAAACGTTATATAAAGACTATGGATCAATAG	1128				
QY	856	TTTAGAAAAAGATGTGATCCGTGAGCGGTTTTCAAAATTTGCAACCGAGAAATGAATTAATA	915				
DB	1129	TTTAGAAAAAGATGTGATCCGTGAGCGGTTTTCAAAATTTGCAACCGAGAAATGAATTAATA	1188				
QY	916	TCCCTTTTATCAAGAGCGCAAAAGAAAAACGAAATGATACCACAATCAGTGCAAAAAAA	975				
DB	1189	TCCCTTTTATCAAGAGCGCAAAAGAAAAACGAAATGATACCACAATCAGTGCAAAAAAA	1248				
QY	976	GATATAATGGGAGATAGACGGTTTCTGTTGCTGCTGACTTGCAACCATATCATAAAAATC	1035				
DB	1249	GATATAATGGGAGATAGACGGTTTCTGTTGCTGCTGACTTGCAACCATATCATAAAAATC	1308				
QY	1036	GAACACGAAAGAAATGCGGAAACGTAAGAAAGAGTTATGAAATTAAGACTTAGAAGCAAA	1095				
DB	1309	GAACACGAAAGAAATGCGGAAACGTAAGAAAGAGTTATGAAATTAAGACTTAGAAGCAAA	1368				
QY	1096	CTTAAGAGTGTGTTGATAGTCAGATATCTTAAAAATTTGTATATAGCAATTCGAAGTTAA	1155				
DB	1369	CTTAAGAGTGTGTTGATAGTCAGATATCTTAAAAATTTGTATATAGCAATTCGAAGTTAA	1428				
QY	1156	ATTAGATGCTAAAAATTTGTAATTAAGAGGAGTGATTAATGAACAAAAATATAAAATA	1215				
DB	1429	ATTAGATGCTAAAAATTTGTAATTAAGAGGAGTGATTAATGAACAAAAATATAAAATA	1488				
QY	1216	TTCTCAAAACTTTTAAACGAGTGAAAAAGTACTCAACCAATTAATAAACAATTTGAATTT	1275				
DB	1489	TTCTCAAAACTTTTAAACGAGTGAAAAAGTACTCAACCAATTAATAAACAATTTGAATTT	1548				
QY	1276	AAAAGAAACCGATACCGTTTACGAAATTTGAAACAGGTAAAGGCAATTTAACGACGAACT	1335				
DB	1549	AAAAGAAACCGATACCGTTTACGAAATTTGAAACAGGTAAAGGCAATTTAACGACGAACT	1608				
QY	1336	GGCTAAATTAAGTAAACAGGTAAAGTATTTGAAATTAGACAGTCACTATTTCAACTTATC	1395				
DB	1609	GGCTAAATTAAGTAAACAGGTAAAGTATTTGAAATTAGACAGTCACTATTTCAACTTATC	1668				
QY	1396	GTCAAGAAAAATTAACCTGAATTAATCTCGTGTCACTTTAAATTCACCAAGATATTTCTACAGTT	1455				
DB	1669	GTCAAGAAAAATTAACCTGAATTAATCTCGTGTCACTTTAAATTCACCAAGATATTTCTACAGTT	1728				
QY	1456	TCAATTTCCCTTAAACAAACAGAGGTATAAAATTTGTTGGAGTATTTCCCTTACATTTAAGCAC	1515				
DB	1729	TCAATTTCCCTTAAACAAACAGAGGTATAAAATTTGTTGGAGTATTTCCCTTACATTTAAGCAC	1788				

Qy	1516	ACAAATTTATTAATAAGAGTGGTTTTTGAAGCCATCGCTGACATCTATCTGATTGTGA	1575
Db	1789	ACAAATTTATTAATAAGAGTGGTTTTTGAAGCCATCGCTGACATCTATCTGATTGTGA	1848
Qy	1576	AGAAGATTTCTACAAGCGTACCTTGGATATTCACCGAACATAGGGTTGCTTTCGACAC	1635
Db	1849	AGAAGATTTCTACAAGCGTACCTTGGATATTCACCGAACATAGGGTTGCTTTCGACAC	1908
Qy	1636	TCAAGTCTCGATTACGAATTCCTTAAGCTGCGAGCAATGCTTTTCATCTCTAAACCAA	1695
Db	1909	TCAAGTCTCGATTACGAATTCCTTAAGCTGCGAGCAATGCTTTTCATCTCTAAACCAA	1968
Qy	1696	AGTAAACAGTGTCTTAATAAACTTACCGCCATACCAAGATGTTCCAGATAAATTTG	1755
Db	1969	AGTAAACAGTGTCTTAATAAACTTACCGCCATACCAAGATGTTCCAGATAAATTTG	2028
Qy	1756	GAACTATATACGTAATTTGTTTCAAAATGGGTCAATCGAGAAATCGTCAACTGTTAC	1815
Db	2029	GAACTATATACGTAATTTGTTTCAAAATGGGTCAATCGAGAAATCGTCAACTGTTAC	2088
Qy	1816	TAAATACAGTTTTCATCAAGCAATGAACAGCCAAAGTAAACAAATTTAAGTACCGTTAC	1875
Db	2089	TAAATACAGTTTTCATCAAGCAATGAACAGCCAAAGTAAACAAATTTAAGTACCGTTAC	2148
Qy	1876	TTATGAGCAAGTATTTGTTTCAAAATGGGTCAATCGAGAAATCGTCAACTGTTAC	1935
Db	2149	TTATGAGCAAGTATTTGTTTCAAAATGGGTCAATCGAGAAATCGTCAACTGTTAC	2208
Qy	1936	CTATGAGTGGCTTTTGAATTTGAAAGTTTACAGTTTACAGTTTAAAGCGGAATGAGATAAT	1995
Db	2209	CTATGAGTGGCTTTTGAATTTGAAAGTTTACAGTTTACAGTTTAAAGCGGAATGAGATAAT	2268
Qy	1996	TTATTAGTATATCTAGCAGCTTCCAAAGAGCTTAAAGAGTCCCTAGCGTCTTTATCAT	2055
Db	2269	TTATTAGTATATCTAGCAGCTTCCAAAGAGCTTAAAGAGTCCCTAGCGTCTTTATCAT	2328
Qy	2056	GGGAGAGTCCGATCATATGCAAGCAAAATAAACTCGCAACAGCACTTGGAGAAATGGG	2115
Db	2329	GGGAGAGTCCGATCATATGCAAGCAAAATAAACTCGCAACAGCACTTGGAGAAATGGG	2388
Qy	2116	ACGAATTCGAGAAAAACCTTTTACGCTGGATATCATATCTAATAAGCGGTAAAGAGACG	2175
Db	2389	ACGAATTCGAGAAAAACCTTTTACGCTGGATATCATATCTAATAAGCGGTAAAGAGACG	2448
Qy	2176	GGTTCAAAAGGTTTAAATTAAGAGAGAGCAATCAATGCATTTAGCTAGAACTATATTTT	2235
Db	2449	GGTTCAAAAGGTTTAAATTAAGAGAGAGCAATCAATGCATTTAGCTAGAACTATATTTT	2508
Qy	2236	TGGACAAAGTGGAGAAATTTAGAGAACGCTCTCCAGACCAAGTTTCAAGAGCTAGTGC	2295
Db	2509	TGGACAAAGTGGAGAAATTTAGAGAACGCTCTCCAGACCAAGTTTCAAGAGCTAGTGC	2568
Qy	2296	ACTAAACATATTTAAGCTTAAGTGTGGAACTGTATATATATGGAAGAAAGCGGT	2355
Db	2569	ACTAAACATATTTAAGCTTAAGTGTGGAACTGTATATATATGGAAGAAAGCGGT	2628
Qy	2356	AGAAATTAAGCAAGAGAGGAAATTTAGAGAGCAATTTAATGCCATATGCTGCGCGTT	2415
Db	2629	AGAAATTAAGCAAGAGAGGAAATTTAGAGAGCAATTTAATGCCATATGCTGCGCGTT	2688
Qy	2416	AGGATGGGAACATATCAATTTTCTGGAGATAAATAATTTCAAGAGATTACACATGG	2475
Db	2689	AGGATGGGAACATATCAATTTTCTGGAGATAAATAATTTCAAGAGATTACACATGG	2748
Qy	2476	GCATAATGAATTTACGCTTCTTACGTTATTAAGAGCGGTTTTTCTTTAATTAAGCGGTC	2535
Db	2749	GCATAATGAATTTACGCTTCTTACGTTATTAAGAGCGGTTTTTCTTTAATTAAGCGGTC	2808
Qy	2536	TTTTTATAGAAAAATCCTTAGCGTGGTTTTTTTTCCGAAATGCTGGCGGTACCCCAAGAA	2595
Db	2809	TTTTTATAGAAAAATCCTTAGCGTGGTTTTTTTTCCGAAATGCTGGCGGTACCCCAAGAA	2868
Qy	2596	TTAGAAATGAGTAGTCAATTTATTCAGAAATAGATCAGAAATCAGATCCAAACATA	2655







QY	1095	ACTTAAGAGTGTGTTGATGCGATCTTAAATTTTGTATAATAGGAATTTGAAGTTA	1154	QY	2175	GGTTTCAAAAAGGTTTAAATTAAGGAGAGCAATCAATGCATTAGCTAGACTATATTTT	2234
DB	1504	ACTTAAGAGTGTGTTGATGCGATCTTAAATTTTGTATAATAGGAATTTGAAGTTA	1563	DB	2584	GGTTTCAAAAAGGTTTAAATTAAGGAGAGCAATCAATGCATTAGCTAGACTATATTTT	2643
QY	1155	AATTAGATGCTTAAAAATTTGTATTAAGAAGGAGTGATTCATGAACAACAAAATATAAAT	1214	QY	2235	TTGGACAACGTTGGAGAAATTTAGAGAACGTCCTCCAGACACAGTTTACAAAGAGCTAGTG	2294
DB	1564	AATTAGATGCTTAAAAATTTGTATTAAGAAGGAGTGATTCATGAACAACAAAATATAAAT	1623	DB	2644	TTGGACAACGTTGGAGAAATTTAGAGAACGTCCTCCAGACACAGTTTACAAAGAGCTAGTG	2703
QY	1215	ATTCTCAAACTTTTAAACGAGTGCAAAAGTACTCAACCAAAATATAAACAATTTGAAT	1274	QY	2295	CACTAAAACATTAATTAACGCTATAAGTGTGTGGAACACACTGTATATATGAAAAAGCCG	2354
DB	1624	ATTCTCAAACTTTTAAACGAGTGCAAAAGTACTCAACCAAAATATAAACAATTTGAAT	1683	DB	2704	CACTAAAACATTAATTAACGCTATAAGTGTGTGGAACACACTGTATATATGAAAAAGCCG	2763
QY	1275	TAAAGAAACCGATACACGTTTACGAAATTTGAAACAGGTAAAGGCACTTTAACGACGAAC	1334	QY	2355	TAGAGAATTTAAACCAAGAGAGAGAAATTTAGAGAAGATTTAATGCCATATCCGTGGCCGT	2414
DB	1684	TAAAGAAACCGATACACGTTTACGAAATTTGAAACAGGTAAAGGCACTTTAACGACGAAC	1743	DB	2764	TAGAGAATTTAAACCAAGAGAGAGAAATTTAGAGAAGATTTAATGCCATATCCGTGGCCGT	2823
QY	1335	TGGCTAAATTAAGTAACAGGTAACGTTTATTTGAATTTAGACAGTCACTATTCAACTTAT	1394	QY	2415	TAGGATGGGAACATATCAATTTTCTTGAGAAATACAAATTTGAAGGATTAATGACACTG	2474
DB	1744	TGGCTAAATTAAGTAACAGGTAACGTTTATTTGAATTTAGACAGTCACTATTCAACTTAT	1803	DB	2824	TAGGATGGGAACATATCAATTTTCTTGAGAAATACAAATTTGAAGGATTAATGACACTG	2883
QY	1395	CGTCAGAAAAATTAACACTGAATCTCGTGTCACTTTAATTCACCAAGATATTTCTACAGT	1454	QY	2475	GCCTAATGAATTTACGTCCTTTACGTATAAAGAGCCGTTTTATTCTTAAATATAACGGCT	2534
DB	1804	CGTCAGAAAAATTAACACTGAATCTCGTGTCACTTTAATTCACCAAGATATTTCTACAGT	1863	DB	2884	GCCTAATGAATTTACGTCCTTTACGTATAAAGAGCCGTTTTATTCTTAAATATAACGGCT	2943
QY	1455	TTCAATCCCTAACAAACAGAGGTATAAATTTGTGGAGTATTCCTTCAATTAAGCA	1514	QY	2535	CTTTTATAGAAAAAATCCTTAGCGTGTGTTTTTCCGAAATGCTGGCGGTACCCCAAGA	2594
DB	1864	TTCAATCCCTAACAAACAGAGGTATAAATTTGTGGAGTATTCCTTCAATTAAGCA	1923	DB	2944	CTTTTATAGAAAAAATCCTTAGCGTGTGTTTTTCCGAAATGCTGGCGGTACCCCAAGA	3003
QY	1515	CACAAATTTAATAAAGAGTGTGTTTGAAGCCATGCTGTGACATCTATCTGATGTTG	1574	QY	2595	ATTAGAAATGAGTAGATCAATTTTACGATAGAAATCAGGMAATCAGATCCAAACCAT	2654
DB	1924	CACAAATTTAATAAAGAGTGTGTTTGAAGCCATGCTGTGACATCTATCTGATGTTG	1983	DB	3004	ATTAGAAATGAGTAGATCAATTTTACGATAGAAATCAGGMAATCAGATCCAAACCAT	3063
QY	1575	AAGAAGGATTTCAAGCGTACCTTGGATATTCACCGAACACTAGGTTGCTTTTGACA	1634	QY	2655	AAAAACACTAGAACAAATGCAAGCTTAACTCAACGCTAGTAGTGGATTTAATCCC	2714
DB	1984	AAGAAGGATTTCAAGCGTACCTTGGATATTCACCGAACACTAGGTTGCTTTTGACA	2043	DB	3064	AAAAACACTAGAACAAATGCAAGCTTAACTCAACGCTAGTAGTGGATTTAATCCC	3123
QY	1635	CTCAAGTCTGATTCAGCAATGCTTAAGCTGCCAGCGGAATGCTTTCATCCTTAAACCA	1694	QY	2715	AAATGAGCCAAACAGAACCCAGAGCCAGAAACAGAAATCAGAACTAGAAATTTGA	2774
DB	2044	CTCAAGTCTGATTCAGCAATGCTTAAGCTGCCAGCGGAATGCTTTCATCCTTAAACCA	2103	DB	3124	AAATGAGCCAAACAGAACCCAGAGCCAGAAACAGAAATCAGAACTAGAAATTTGA	3183
QY	1695	AAGTAAACAGTCTCTTAATAAACTTACCCGCCATACCAACAGATGTTTCCAGATAAATTT	1754	QY	2775	AATGGAAGAGAAAAAAGCAATGCTTGTGTAATTAATGCAAGAAATCGTTGCTTATTT	2834
DB	2104	AAGTAAACAGTCTCTTAATAAACTTACCCGCCATACCAACAGATGTTTCCAGATAAATTT	2163	DB	3184	AATGGAAGAGAAAAAAGCAATGCTTGTGTAATTAATGCAAGAAATCGTTGCTTATTT	3243
QY	1755	GGAAGCTATACGTAATGTTTCAAAATGGTCAATCGAGATATCGTCAACTGTTTA	1814	QY	2835	TTTTTTAAAGCGGTATACCTAGATATAACGAAACCAACGAACTGAAATAGAAACGAA	2894
DB	2164	GGAAGCTATACGTAATGTTTCAAAATGGTCAATCGAGATATCGTCAACTGTTTA	2223	DB	3244	TTTTTTAAAGCGGTATACCTAGATATAACGAAACCAACGAACTGAAATAGAAACGAA	3303
QY	1815	CTAAAAACAGTTTCATCAAGCAATGAACACCGCCAAAGTAAACAAATTTAAGTACCGTTA	1874	QY	2895	AGCCATGACACATTTTAAATGTTTGAAGCAATTTTAAATGATAGCCCGGATAAGAT	2954
DB	2224	CTAAAAACAGTTTCATCAAGCAATGAACACCGCCAAAGTAAACAAATTTAAGTACCGTTA	2283	DB	3304	AGCCATGACACATTTTAAATGTTTGAAGCAATTTTAAATGATAGCCCGGATAAGAT	3363
QY	1875	CTTATGAGCAAGTATGCTATTTTAAATGATTAATTTTAAACGGGAGGAATTAAT	1934	QY	2955	TGCAAAACCAACGCTTATCAGTTAGTTCAGATGAACCTTTCCCTCGTAAGAGATTTTAA	3014
DB	2284	CTTATGAGCAAGTATGCTATTTTAAATGATTAATTTTAAACGGGAGGAATTAAT	2343	DB	3364	TGCAAAACCAACGCTTATCAGTTAGTTCAGATGAACCTTTCCCTCGTAAGAGATTTTAA	3423
QY	1935	TCTATGAGTCGCTTTGTAAATTTGGAAGTTTACACGTTTCTAAGAGGAATGTAGATAAA	1994	QY	3015	TTAACTTTGTTTGAAGACCGGTATTAACCGTACTATTAACCGTACTATTAAGGAAATCAGAGAGT	3074
DB	2344	TCTATGAGTCGCTTTGTAAATTTGGAAGTTTACACGTTTCTAAGAGGAATGTAGATAAA	2403	DB	3424	TTAACTTTGTTTGAAGACCGGTATTAACCGTACTATTAACCGTACTATTAAGGAAATCAGAGAGT	3483
QY	1995	TTATTAGGTATACCTGACAGCTTCCAGAGAGCTTAAAGAGGTCCTTAGCGCTCTTATCA	2054	QY	3075	TTTCAAGTATCTAAGCTACTGAAATTTAAGAAATTTGTTAAGCAATCAATCGGAATCGTTTG	3134
DB	2404	TTATTAGGTATACCTGACAGCTTCCAGAGAGCTTAAAGAGGTCCTTAGCGCTCTTATCA	2463	DB	3484	TTTCAAGTATCTAAGCTACTGAAATTTAAGAAATTTGTTAAGCAATCAATCGGAATCGTTTG	3543
QY	2055	TGGGAGAGCTCGGATCATATGCAAGCAAAATAAATCGCAACAGCACTTGGAGAAATGG	2114	QY	3135	ATTGCTTTTTTGTATTCATTTATAGAGGTGGAGTTTGTATGAATCATGATGAATGA	3194
DB	2464	TGGGAGAGCTCGGATCATATGCAAGCAAAATAAATCGCAACAGCACTTGGAGAAATGG	2523	DB	3544	ATTGCTTTTTTGTATTCATTTATAGAGGTGGAGTTTGTATGAATCATGATGAATGA	3603
QY	2115	GACGAATCGAGAAACCCCTCTTTAGCGTGGATTACATATCTAATAAGCCGTAAGAGAC	2174	QY	3195	AACTTATATAAAAAATAGTTTTATTGGAGATAGAAAAATTTAGCAAAATATCTATACACTAG	3254
DB	2524	GACGAATCGAGAAACCCCTCTTTAGCGTGGATTACATATCTAATAAGCCGTAAGAGAC	2583	DB	3604	AACTTATATAAAAAATAGTTTTATTGGAGATAGAAAAATTTAGCAAAATATCTATACACTAG	3663
				QY	3255	AACGTTTAAAGAAAGGTTTAGAAAAAGAAATATCTACTTTAGAAACAAAAATCAGATAAGTA	3314



Db 3664 AACGTTTAAAGAAAGAGTTAGAAAAGAGAAATATCTACTAGAAAACAAAATCAGATAAGTA 3723  
QY TTTTCTTCGAGGGGGAAGATTATATATATAGCTTAATAGAAAATAACAAATATATTTA 3374  
Db 3724 TTTTCTTCGAGGGGGAAGATTATATATATAGCTTAATAGAAAATAACAAATATATTTA 3783  
QY TTCGATTAGTGGAAAAAAATGACTTATAAGGAAAAAAATCTTTTCAAAACATGCAAT 3434  
Db 3784 TTCGATTAGTGGAAAAAAATGACTTATAAGGAAAAAAATCTTTTCAAAACATGCAAT 3843  
QY ATTCAAAACAGTTGAATCAAAAAGCAAAACCAAGTTTAATTAACACCTATTATTAGATT 3494  
Db 3844 ATTGAACAGTTGAATGAAAAGCAAAACCAAGTTTAATTAACACCTATTATTAGATT 3903  
QY TATAGAAAAGAGAACAGCTGAATGAATATCCCTTTTGTGTAGAAAATCTGTCTCATGA 3554  
Db 3904 TATAGAAAAGAGAACAGCTGAATGAATATCCCTTTTGTGTAGAAAATCTGTCTCATGA 3963  
QY CGGCTTCGTTAAAGTACAAAATTTAAAATAGTAAAATTCGCTCAATCACTACCAAGCCAGG 3614  
Db 3964 CGGCTTCGTTAAAGTACAAAATTTAAAATAGTAAAATTCGCTCAATCACTACCAAGCCAGG 4023  
QY TAAAGCAAAAGGGGCTATTTTGGCTATCGCTCAAAATCAAGCATGATTGGCGGTCGTGG 3674  
Db 4024 TAAAGCAAAAGGGGCTATTTTGGCTATCGCTCAAAATCAAGCATGATTGGCGGTCGTGG 4083  
QY TGTGTTCTGACTTCCAGGAAGCGATTCAAGAAAATCAAGATACATTACACATTGGAC 3734  
Db 4084 TGTGTTCTGACTTCCAGGAAGCGATTCAAGAAAATCAAGATACATTACACATTGGAC 4143  
QY ACCCAAGCTTTATCGTTATGGAAGCTATGCAAGCAAAACCGTTTCATACAGAAAGGACA 3794  
Db 4144 ACCCAAGCTTTATCGTTATGGAAGCTATGCAAGCAAAACCGTTTCATACAGAAAGGACA 4203  
QY TCTGAAAAACAAATTAAAGCAAAATCAATACCTTTCTTTATGATTTTGAATTCACACGGC 3854  
Db 4204 TCTGAAAAACAAATTAAAGCAAAATCAATACCTTTCTTTATGATTTTGAATTCACACGGC 4263  
QY AAAAGAAACCTATTCAGCAGCGCATATTTAAACCGCTATTGATTTAGTTTATGCG 3914  
Db 4264 AAAAGAAACCTATTCAGCAGCGCATATTTAAACCGCTATTGATTTAGTTTATGCG 4323  
QY TACTATGATTATCAAAATCTGATAAAGGTTATCAAGCATATTTTGTGTTTAAAGCCGCT 3974  
Db 4324 TACTATGATTATCAAAATCTGATAAAGGTTATCAAGCATATTTTGTGTTTAAAGCCGCT 4383  
QY CTATGCTCAAAATCAAGATTTAAATCTGTAAGCAAGCAAGCAAAATTTTCGCAAAA 4034  
Db 4384 CTATGCTCAAAATCAAGATTTAAATCTGTAAGCAAGCAAGCAAAATTTTCGCAAAA 4443  
QY TATCCGAGATATTTTGGAAAGTCTTTGCCAGTTGATCTAACGTGTAATCATTTTCGTAT 4094  
Db 4444 TATCCGAGATATTTTGGAAAGTCTTTGCCAGTTGATCTAACGTGTAATCATTTTCGTAT 4503  
QY TGCTCGCATACCAAGAACCGCAATGTAGAAATTTTGTGATCTCTTAATTCGTTATCTTT 4154  
Db 4504 TGCTCGCATACCAAGAACCGCAATGTAGAAATTTTGTGATCTCTTAATTCGTTATCTTT 4563  
QY CAAAGAAATGCAAGATTGCTTTTCAAAACCAACAGATTAATAGGGCTTTTACTCGTTCAAG 4214  
Db 4564 CAAAGAAATGCAAGATTGCTTTTCAAAACCAACAGATTAATAGGGCTTTTACTCGTTCAAG 4623  
QY TCTAACCGTTTAAAGCGGTACAGAGGCAAAACCAAGTAGATGAACCTGTTTAAATCT 4274  
Db 4624 TCTAACCGTTTAAAGCGGTACAGAGGCAAAACCAAGTAGATGAACCTGTTTAAATCT 4683  
QY CTTATTGCAAGAAACGAATTTTCAGAGAAAGGGTTTAAATAGGCGGTAAATAACGTCT 4334  
Db 4684 CTTATTGCAAGAAACGAATTTTCAGAGAAAGGGTTTAAATAGGCGGTAAATAACGTCT 4743  
QY GTTTACCTCTCTTTAGCCCTACTTTTAGTTTCAAGCTATTCOAATCGAAACGTCGGAATATA 4394

4744 GTTTACCTCTCTTTAGCCCTACTTTTAGTTTCAGGCTATTCAATCGAAACGTCGGAATATA 4803  
QY TATGTTGAGTTTAAATATCGATTAGTCAACCCCTTAGAGAAAAGAAAGTAAATCAAAAT 4454  
Db 4804 TATGTTGAGTTTAAATATCGATTAGTCAACCCCTTAGAGAAAAGAAAGTAAATCAAAAT 4863  
QY TGTTAGAAGTGCCTATTTCAGAAAACCTATCAAGGGGCTTAATAGGGAATACATTACCATTCT 4514  
Db 4864 TGTTAGAAGTGCCTATTTCAGAAAACCTATCAAGGGGCTTAATAGGGAATACATTACCATTCT 4923  
QY TTGCAAGCTTCGGGTATCAAGTGATTAAACCAAGTAAAGATTATTTGTCCTCAAGGGTG 4574  
Db 4924 TTGCAAGCTTCGGGTATCAAGTGATTAAACCAAGTAAAGATTATTTGTCCTCAAGGGTG 4983  
QY GTTTAAATTCAGAAAAAAGAAAGCAAGCTCAACGCTGTTTCATTGTCAGAAATGAAAGA 4634  
Db 4984 GTTTAAATTCAGAAAAAAGAAAGCAAGCTCAACGCTGTTTCATTGTCAGAAATGAAAGA 5043  
QY AGATTTAATGCTTATATTAGGCAAAAAGCAAGTATACAAAGCCTTATTTAGTGACGAC 4694  
Db 5044 AGATTTAATGCTTATATTAGGCAAAAAGCAAGTATACAAAGCCTTATTTAGTGACGAC 5103  
QY CAAAAAGAGATTAGAGAAAGTCTAGGCAATTCCTGAACGGACATTAGATAAATTCCTGAA 4754  
Db 5104 CAAAAAGAGATTAGAGAAAGTCTAGGCAATTCCTGAACGGACATTAGATAAATTCCTGAA 5163  
QY GGTACTGAAGGCGAATTCAGAAAATTTTCTTAAGATTAAACCAAGAAATATGTTGGCAT 4814  
Db 5164 GGTACTGAAGGCGAATTCAGAAAATTTTCTTAAGATTAAACCAAGAAATATGTTGGCAT 5223  
QY TCAACTGCTAGTGTAAATCATTTGCTCATTCGATCATTAAAGTAAAGAAAGAAAGAAAA 4874  
Db 5224 TCAACTGCTAGTGTAAATCATTTGCTCATTCGATCATTAAAGTAAAGAAAGAAAGAAAA 5283  
QY AGAAAGCTATATAAAGGCGCTGACAAAATTTCTTTGACTTAGAGCATACATTCAATCAAGA 4934  
Db 5284 AGAAAGCTATATAAAGGCGCTGACAAAATTTCTTTGACTTAGAGCATACATTCAATCAAGA 5343  
QY GACTTTAAACAGCTAGCAGAACCGCTTAAACCGCACACACAACTCGATTTGTTTAGCTA 4994  
Db 5344 GACTTTAAACAGCTAGCAGAACCGCTTAAACCGCACACACAACTCGATTTGTTTAGCTA 5403  
QY TGATACAGCTGAAAATTAACCCCGCACATATGCCATTACATTTATCTATGATACGTTG 5054  
Db 5404 TGATACAGCTGAAAATTAACCCCGCACATATGCCATTACATTTATCTATGATACGTTG 5463  
QY TTGTTTTTTCTTGTGTTTACGGAATGATTAGCAGAAAATATACAGAGTAAAGTTTAAAT 5114  
Db 5464 TTGTTTTTTCTTGTGTTTACGGAATGATTAGCAGAAAATATACAGAGTAAAGTTTAAAT 5523  
QY TAAATTATTAGGGGAGAGAGAGTAGCCCGAAAACCTTTTAGTTGCTTGGACTGAAC 5174  
Db 5524 TAAATTATTAGGGGAGAGAGAGTAGCCCGAAAACCTTTTAGTTGCTTGGACTGAAC 5583  
QY GAAAGTGGGAAAGGCTACTAAAAACGTCGAGGGGCGAGTAGAGCGAAACGCAACCTTGAT 5234  
Db 5584 GAAAGTGGGAAAGGCTACTAAAAACGTCGAGGGGCGAGTAGAGCGAAACGCAACCTTGAT 5643  
QY TTTTAAATTTCTCTTTTATAGGTCATTAGAGTATACATTATTTGCTTATTAACCTATT 5294  
Db 5644 TTTTAAATTTCTCTTTTATAGGTCATTAGAGTATACATTATTTGCTTATTAACCTATT 5703  
QY TAGCAGCATATATAGATTATTGAAATAGGTCATTATTAAGTTGAGCATATTAGAGGAGGAAAA 5354  
Db 5704 TAGCAGCATATATAGATTATTGAAATAGGTCATTATTAAGTTGAGCATATTAGAGGAGGAAAA 5763  
QY TCTTGGAGAAATATTTCAGAACCCGATTACATGGAATGGAATAGTTCTTGTGTTAGCT 5414  
Db 5764 TCTTGGAGAAATATTTCAGAACCCGATTACATGGAATGGAATAGTTCTTGTGTTAGCT 5823  
QY GGTTTTTAACTAAAGTAGTGAATTTTGAATTTTGTGTTGTTGTTGTTGTTGTTGTTGTT 5474  
Db 5824 GGTTTTTAACTAAAGTAGTGAATTTTGAATTTTGTGTTGTTGTTGTTGTTGTTGTTGTT 5883



QY 1759 GCTATATACGTACTTTGTTTCAAAATCGGTCAATCGAGATATCGTCAACTGTTTACTAA 1818  
DB 2132 GCTATATACGTACTTTGTTTCAAAATCGGTCAATCGAGATATCGTCAACTGTTTACTAA 2191  
QY 1819 AAATCAGTTTCATCAACCAATGAACACGCGCAAGTAAACAAATTAAGTACCGTTACTTA 1878  
DB 2192 AAATCAGTTTCATCAACCAATGAACACGCGCAAGTAAACAAATTAAGTACCGTTACTTA 2251  
QY 1879 TGACCAAGTATGTCATATTTTAAATAGTTATCTATATTTAAATTAAGGAGGAAATTAATCTTA 1938  
DB 2252 TGACCAAGTATGTCATATTTTAAATAGTTATCTATATTTAAATTAAGGAGGAAATTAATCTTA 2311  
QY 1939 TGAGTCGCTTTGTTGAAATTTGGAAGTTTACAGTTTACTTAAGGGAATGTAGATAAATTAAT 1998  
DB 2312 TGAGTCGCTTTGTTGAAATTTGGAAGTTTACAGTTTACTTAAGGGAATGTAGATAAATTAAT 2371  
QY 1999 TAGGTATATCTGACAGCTTCCAAAGAGCTAAAGAGTCCCTAGCGCTCTTATCATGGG 2058  
DB 2372 TAGGTATATCTGACAGCTTCCAAAGAGCTAAAGAGTCCCTAGCGCTCTTATCATGGG 2431  
QY 2059 GAAGCTCGGATCATATCGACAGCAAAATAAATTAAGTCCGCAACAGCACTTGGAGAAATGGGAG 2118  
DB 2432 GAAGCTCGGATCATATCGACAGCAAAATAAATTAAGTCCGCAACAGCACTTGGAGAAATGGGAG 2491  
QY 2119 AATCGAGAAAACCTCTTTACGCTGGATTACATATCTAATAAGCCGTAAGGAGACGGGT 2178  
DB 2492 AATCGAGAAAACCTCTTTACGCTGGATTACATATCTAATAAGCCGTAAGGAGACGGGT 2551  
QY 2179 TCAAAAAGGTTTAAATAAGGAGAGCAATCAATGCAATAGCTAGTAACTATATTTTGG 2238  
DB 2552 TCAAAAAGGTTTAAATAAGGAGAGCAATCAATGCAATAGCTAGTAACTATATTTTGG 2611  
QY 2239 ACAAGCTGGAGAAATTTAGAGAACTGCTCTCCAAAGCCAGTTTCAAGAGCTAGTGCAC 2298  
DB 2612 ACAAGCTGGAGAAATTTAGAGAACTGCTCTCCAAAGCCAGTTTCAAGAGCTAGTGCAC 2671  
QY 2299 AAACATAATTTAAACGCTATAAGTGTGCGAACACTGTATATATGGAAGAAACCGTAG 2358  
DB 2672 AAACATAATTTAAACGCTATAAGTGTGCGAACACTGTATATATGGAAGAAACCGTAG 2731  
QY 2359 AGAATTAAGAGAGAGAGGAAATTTAGAGAACTTAAATGCAATGCTAGTGGCGGTAGG 2418  
DB 2732 AGAATTAAGAGAGAGGAAATTTAGAGAACTTAAATGCAATGCTAGTGGCGGTAGG 2791  
QY 2419 ATGGGAACATATCAATTTTCTTGAGAAATCAAAATTTGAAGGATTAATGACACTGGGA 2478  
DB 2792 ATGGGAACATATCAATTTTCTTGAGAAATCAAAATTTGAAGGATTAATGACACTGGGA 2851  
QY 2479 AATGAATTTAGCTCTTTAGCTATAAAGAGCGGTTTATCTTAATATAACGCGCTCTTT 2538  
DB 2852 AATGAATTTAGCTCTTTAGCTATAAAGAGCGGTTTATCTTAATATAACGCGCTCTTT 2911  
QY 2539 TTATAGAAAAATCCTTTAGCGTGGTTTTTTTCCGAAATGCTGGCGGTACCCCAAGAAATTA 2598  
DB 2912 TTATAGAAAAATCCTTTAGCGTGGTTTTTTTCCGAAATGCTGGCGGTACCCCAAGAAATTA 2971  
QY 2599 GAAATAGTATGATCAAAATTTTCAAGAAATGATCAAGAAATCAGAGAAATCAGATCAAA 2658  
DB 2972 GAAATAGTATGATCAAAATTTTCAAGAAATGATCAAGAAATCAGAGAAATCAGATCAAA 3031  
QY 2659 ACATAGAAACAAATTTGCAAGTAACTAATCAAGCTAGTGTAGTTTAAATCCCAAT 2718  
DB 3032 ACATAGAAACAAATTTGCAAGTAACTAATCAAGCTAGTGTAGTTTAAATCCCAAT 3091  
QY 2719 GAGCAACAGAACCCAGAGCCAGAAACAGAAATCAGAACAGTAAACATTTGAGAAATG 2778  
DB 3092 GAGCAACAGAACCCAGAGCCAGAAACAGAAATCAGAACAGTAAACATTTGAGAAATG 3151  
QY 2779 GAAGAGAAAAAGCAATGCTTGGTGTGAATTAATGCAAGAAATCGTGTCTTATTTTTT 2838  
DB 3152 GAAGAGAAAAAGCAATGCTTGGTGTGAATTAATGCAAGAAATCGTGTCTTATTTTTT 3211

QY 2839 TTAAAAGCGGTACTAGATATACGAAACAAACGAACTGAATAGAAACGAAAAAGAGCC 2898  
DB 3212 TTAAAAGCGGTACTAGATATACGAAACAAACGAACTGAATAGAAACGAAAAAGAGCC 3271  
QY 2899 ATGACACATTTTATAAAATGTTTGACGACATTTTAAATGATAGCCGATAGATTTGCC 2958  
DB 3272 ATGACACATTTTATAAAATGTTTGACGACATTTTAAATGATAGCCGATAGATTTGCC 3331  
QY 2959 AAACCAACGCTTATCAGTTAGTCAGATGAATCTCTTCCCTCGTAAAGATTTAATTA 3018  
DB 3332 AAACCAACGCTTATCAGTTAGTCAGATGAATCTCTTCCCTCGTAAAGATTTAATTA 3391  
QY 3019 CTTTGTGTTGAAGAGCGGTATATAACGCTACTATCATATATAGGGAAATCAGAGATTTTC 3078  
DB 3392 CTTTGTGTTGAAGAGCGGTATATAACGCTACTATCATATATAGGGAAATCAGAGATTTTC 3451  
QY 3079 AAGTATCTAAGCTACTGAATTTAAGAAATGTTTAAGCAATCAATCGAAATCGTTGATG 3138  
DB 3452 AAGTATCTAAGCTACTGAATTTAAGAAATGTTTAAGCAATCAATCGAAATCGTTGATG 3511  
QY 3139 CTTTGTGTTGATTCATTTATAGAGGTGGAGTTGTATGAATCATGATGAATTAAT 3198  
DB 3512 CTTTGTGTTGATTCATTTATAGAGGTGGAGTTGTATGAATCATGATGAATTAAT 3571  
QY 3199 TATATAAAAAATAGTTTATGAGATAAGAAAAATTAGCAATATCTATACACTAGAAACG 3258  
DB 3572 TATATAAAAAATAGTTTATGAGATAAGAAAAATTAGCAATATCTATACACTAGAAACG 3631  
QY 3259 TTTTAAAGAGGTTAGAAAAAGAAATATCTACTTTAGAAAAAATAATCAGATAAGTATTT 3318  
DB 3632 TTTTAAAGAGGTTAGAAAAAGAAATATCTACTTTAGAAAAAATAATCAGATAAGTATTT 3691  
QY 3319 TCTTCGGAGGGGAGAGATTATATATATAGTTTAAATAGAAAAATAACAAAAATTTTATTCG 3378  
DB 3692 TCTTCGGAGGGGAGAGATTATATATATAGTTTAAAGAAAAATACTTTTCAAAAATGCAATTCG 3751  
QY 3379 ATTAGTGAATAAAAAATGACTTATTAAGAAAAATACTTTTCAAAAATGCAATTCG 3438  
DB 3752 ATTAGTGAATAAAAAATGACTTATTAAGAAAAATACTTTTCAAAAATGCAATTCG 3811  
QY 3439 AAACAGTTGATTAAGAAAAAGCAACCAAGTTAATAAACAACCTATTTTATAGGATTTATA 3498  
DB 3812 AAACAGTTGATTAAGAAAAAGCAACCAAGTTAATAAACAACCTATTTTATAGGATTTATA 3871  
QY 3499 GGAAGAGGAGACAGCTGAATATCCCTTTGTTGTAGAAAACTGTGCTTCATGACGCG 3558  
DB 3872 GGAAGAGGAGACAGCTGAATATCCCTTTGTTGTAGAAAACTGTGCTTCATGACGCG 3931  
QY 3559 TTGTTTAAAGTACAAATTTTAAATAGTAAATTCGCTCAATCACTACCAAGCCAGTAAA 3618  
DB 3932 TTGTTTAAAGTACAAATTTTAAATAGTAAATTCGCTCAATCACTACCAAGCCAGTAAA 3991  
QY 3619 AGCAAGGGGCTATTTTTCGCTATCGCTCAAAATCAAGCATGATGCGGCTCGTGTGT 3678  
DB 3992 AGCAAGGGGCTATTTTTCGCTATCGCTCAAAATCAAGCATGATGCGGCTCGTGTGT 4051  
QY 3679 GTTCTGACTTCGAGGAGAGCGATTCAAGAAATCAAGATACTTTTACATTTGACACACC 3738  
DB 4052 GTTCTGACTTCGAGGAGAGCGATTCAAGAAATCAAGATACTTTTACATTTGACACACC 4111  
QY 3739 AACGTTTATCGTTATGGAACGTTATGACAGCAACCGTTCATACACGAAAGGACATCT 3798  
DB 4112 AACGTTTATCGTTATGGAACGTTATGACAGCAACCGTTCATACACGAAAGGACATCT 4171  
QY 3799 GAAAAATTTTAAAGCAAAATCAATACCTTCTTTATGATTTGATTTTACACCGGCAAAA 3858  
DB 4172 GAAAAATTTTAAAGCAAAATCAATACCTTCTTTATGATTTGATTTTACACCGGCAAAA 4231  
QY 3859 GAACTATTTTACGCAAGCGATTTTAAACACCGCTATTTAGTTTATGCTTACT 3918  
DB 4232 GAACTATTTTACGCAAGCGATTTTAAACACCGCTATTTAGTTTATGCTTACT 4291  
QY 3919 ATGATTATCAAAATCTGATAAAGGTTATCAAGCATATTTTGTGTTAGAAACCCAGTCTAT 3978

4292	Db	ATGATTATCAAACTCTGATAAAGGTTATCAAGCATATTTTGTGTTTAGAAAAGCCAGTCTAT	4351
3979	QY	GTGACTTCAAAATCAAGATTTAAATCTGTCAAGCAGCCAAATATATTTTCGAAAATATC	4038
4352	Db	GTGACTTCAAAATCGAATTTAAATCTGTCAAGCAGCCAAATATATTTTCGAAAATATC	4411
4039	QY	CGAAGATATTTTCGAAAGCTTTGCCAGTTGATCTAAACGTTGAATCATTTTGGTATTGCT	4098
4412	Db	CGAAGATATTTTCGAAAGCTTTGCCAGTTGATCTAAACGTTGAATCATTTTGGTATTGCT	4471
4099	QY	CGCATACCAAGAACGGACAATGTAGAAATTTTTTTTGATCTCTAATTTACCGTTTATCTTTCAA	4158
4472	Db	CGCATACCAAGAACGGACAATGTAGAAATTTTTTTTGATCTCTAATTTACCGTTTATCTTTCAA	4531
4159	QY	GAATGCGAAGATTGGTCTTTTCAAAACAAACAGATAATAAGGGCTTTACTCGTCAAGTCTA	4218
4532	Db	GAATGCGAAGATTGGTCTTTTCAAAACAAACAGATAATAAGGGCTTTACTCGTCAAGTCTA	4591
4219	QY	ACGGTTTTAAGCGGTACAGAGCGCAAAAACAAGTAGATGAACCCCTGGTTTTAATCTCTTTA	4278
4592	Db	ACGGTTTTAAGCGGTACAGAGCGCAAAAACAAGTAGATGAACCCCTGGTTTTAATCTCTTTA	4651
4279	QY	TTGCACGAACGAAATTTTTCAGGAGAAAAAGGGTTTAATAGGGCGTAAATAACGTCATGTTT	4338
4652	Db	TTGCACGAACGAAATTTTTCAGGAGAAAAAGGGTTTAATAGGGCGTAAATAACGTCATGTTT	4711
4339	QY	ACCTCTCTTTAGCCTACTTTTAGTTTCAGGCTATTCAATCGAAACGTCGCAATATAATATG	4398
4712	Db	ACCTCTCTTTAGCCTACTTTTAGTTTCAGGCTATTCAATCGAAACGTCGCAATATAATATG	4771
4399	QY	TTTGAGTTTAATTAATCGATTAGATCAACCTTAGAAGAAAAAGAGTAAATCAAAAATGTT	4458
4772	Db	TTTGAGTTTAATTAATCGATTAGATCAACCTTAGAAGAAAAAGAGTAAATCAAAAATGTT	4831
4459	QY	AGAAGTGCCTTATTCAGAAAACTCAAGGGCTTAATAGGGAATACATTACCATCTTTGC	4518
4832	Db	AGAAGTGCCTTATTCAGAAAACTCAAGGGCTTAATAGGGAATACATTACCATCTTTGC	4891
4519	QY	AAAGCTTTGGTATCAAGTGATTTTAAACGATTAAGATTTATTGTCGCTCAAGGGTGGTTT	4578
4892	Db	AAAGCTTTGGTATCAAGTGATTTTAAACGATTAAGATTTATTGTCGCTCAAGGGTGGTTT	4951
4579	QY	AAATTCAAGAAAAAAGAAGCGAACGTCAACGTGTTTCATTGTCAGAAATGGAAGAAGAT	4638
4952	Db	AAATTCAAGAAAAAAGAAGCGAACGTCAACGTGTTTCATTGTCAGAAATGGAAGAAGAT	5011
4639	QY	TTAATGGCTTATATTAGCGAAAAAAGCGATGTATACAAGCTTATTATTAGTGACGACCAA	4698
5012	Db	TTAATGGCTTATATTAGCGAAAAAAGCGATGTATACAAGCTTATTATTAGTGACGACCAA	5071
4699	QY	AAAGAGATTAGAGAAGTGTAGCATTTCTCGAACGGACATTTAGATAAAATGCTGAAAGTTA	4758
5072	Db	AAAGAGATTAGAGAAGTGTAGCATTTCTCGAACGGACATTTAGATAAAATGCTGAAAGTTA	5131
4759	QY	CTGAAGCGCAATCAGGAAATTTTCTTTAAGATTAAACCGAGGAAGAAATGGTGCAATCAA	4818
5132	Db	CTGAAGCGCAATCAGGAAATTTTCTTTAAGATTAAACCGAGGAAGAAATGGTGCAATCAA	5191
4819	QY	CTTGCTAGTCTTAATTCATTGTTGCTATCCGATCATTTAAAGCTAAAAAGAAAAAGAA	4878
5192	Db	CTTGCTAGTCTTAATTCATTGTTGCTATCCGATCATTTAAAGCTAAAAAGAAAAAGAA	5251
4879	QY	AGCTATATAAGGCGCTGACAAATCTTTTGACTTTAGAGCATACATTCATTCAAGAGACT	4938
5252	Db	AGCTATATAAGGCGCTGACAAATCTTTTGACTTTAGAGCATACATTCATTCAAGAGACT	5311
4939	QY	TTTAAACGAGCTAGCAGNACCGCCCTTAAACGGAACACACTCCGATTTGTTTACGTATGAT	4998
5312	Db	TTTAAACGAGCTAGCAGNACCGCCCTTAAACGGAACACACTCCGATTTGTTTACGTATGAT	5371
4999	QY	ACAGGCTGAAAAATAAAACCGCGCACTATGCCATTTACATTTATATCTATGATACGTGTTGT	5058

5372	ACAGCGTGAATAATAAAACCGCACACTATGCCATTACATTATATATCATATGATACGCTGTTGCT	54331
5059	TTTTTCTTTTGGCTGTTTATAGCGAATGATTAGCAGAAATATACAGAGTAAGATTTTAAATTAAT	5118
5432	TTTTTCTTTTGGCTGTTTATAGCGAATGATTAGCAGAAATATACAGAGTAAGATTTTAAATTAAT	5491
5119	TATTAGGGGGAGAGAGAGAGAGTAGCCCGAAAACCTTTTAGTTGGCTTGGACTGACGGAAG	5178
5492	TATTAGGGGGAGAGAGAGAGTAGCCCGAAAACCTTTTAGTTGGCTTGGACTGACGGAAG	5551
5179	TGAGGGAAGGCTACTATAAAACGTGAGGGGCGCAGTGAGAGCGGAAGCGAAACACTTGAATTTT	5238
5552	TGAGGGAAGGCTACTATAAAACGTGAGGGGCGCAGTGAGAGCGGAAGCGAAACACTTGAATTTT	5611
5239	TAAATTTTCTATCTTTTATATAGGTCACTTAGAGTATATCTTATTGTCTCTATATAAATCTATTAGC	5298
5612	TAAATTTTCTATCTTTTATATAGGTCACTTAGAGTATATCTTATTGTCTCTATATAAATCTATTAGC	5671
5299	AGCATAAATAGATTTATTGAAATAGGTCAATTTAAGATTGAGCATATTAGAGAGGAGAAATCTT	5358
5672	AGCATAAATAGATTTATTGAAATAGGTCAATTTAAGATTGAGCATATTAGAGAGGAGAAATCTT	5731
5359	GGAGAAATATTTGAAGAAACCGAATTACATGGATTGGATTAGTTCTTGTGGTTACGTGGTT	5418
5732	GGAGAAATATTTGAAGAAACCGAATTACATGGATTGGATTAGTTCTTGTGGTTACGTGGTT	5791
5419	TTTTAATCTAAAGTAGTGAATTTTGAATTTTGGTGTGTGTCTTGTCTTGTCTTAGTATTTCG	5478
5792	TTTTAATCTAAAGTAGTGAATTTTGAATTTTGGTGTGTGTCTTGTCTTGTCTTAGTATTTCG	5851
5479	TAGTCAAAGTGATTTAAATA	5497
5852	TAGTCAAAGTGATTTAAATA	5870

## RESULT 5

RESULT 3  
AAX07549  
ID AAX07549 standard: DNA: 5230 BP.

XX  
AC AAX07549:

08-JUN-1999 (first entry)

DE mutant beta-toxin expression cassette.

xx  
kw Beta-toxin; cpb; vaccine; infection; protection; retained immunogenicity;  
kw haemorrhagic enteritis; necrotic enteritis; enterotoxemia;  
kw lamb dysentery; ds.

XX 05 Synthetic.

XX  
PN  
EP892054-A1XX  
20-TAN-1999XX  
DE 17-TTN-1000. 00EP-00202032XX  
PP 20-TTN-1997. 9780-00201888XX  
PA  
ALVTY \ KYZO NOBEL. NYXX  
PT  
**Carcare** PDMM Waterfield NP Frandsen Pl.: Wells JM:

XX  
DP WPT: 1999-083571/08

XX New detoxified derivative of Clostridium perfringens beta-toxin - with  
PT retained immunogenicity, useful as a vaccine to protect against  
PT Clostridium perfringens infection.  
PT

XX  
PS Example: Fig 1c: 30pp: English.

XX  
CC The sequence is that of the pTREX1 expression cassette which was used in  
CC the construction of a detoxified immunogenic derivative of Clostridium  
CC

perfringens beta-toxin. Such a derivative is useful as a vaccine to induce an immune response and protect against Clostridium perfringens infection in man, pig, lamb, sheep, goat, calf and bird, which causes haemorrhagic enteritis, necrotic enteritis, enterotoxemia and lamb dysenteria. The new beta-toxin has been detoxified without impairing the immunogenicity, unlike prior art vaccines which used chemical and formalin detoxifying methods. Production of the beta-toxin in gram positive bacteria prevents prior art difficulties of isolation from dangerous Clostridium perfringens, and purification from the Clostridium perfringens spores

XX Sequence 5230 BP; 1930 A; 795 C; 983 G; 1522 T; 0 U; 0 Other;

Query Match	90.8%;	Score 4993.4;	DB 2;	Length 5230;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 4997;	Conservative	0;	Mismatches 6;	Indels 0;
Gaps	0;			
QY 495	TCCTAACTAGTCCGGTCTTAACAAAGCCCGAAGGAGCTGAGTTGGCTGTGCCA	554		
DB 228	TCCTAGAAATGGATCCGGTCTTAACAAAGCCCGAAGGAGCTGAGTTGGCTGTGCCA	287		
QY 555	CCGCTGAGCAATAACTAGCATACCCCTTGGGGCCCTTAAACGGGCTCTTGAAGGGTCTTT	614		
DB 288	CCGCTGAGCAATAACTAGCATACCCCTTGGGGCCCTTAAACGGGCTCTTGAAGGGTCTTT	347		
QY 615	TGCTGAAAGGAGCACTATATCCGGATGACCTGCGAGCGAAGCTCTAGAAATCGATACGATT	674		
DB 348	TGCTGAAAGGAGCACTATATCCGGATGACCTGCGAGCGAAGCTCTAGAAATCGATACGATT	407		
QY 675	TTCAAGTGGCAACAGATATAAAGGAGCTTTAAATTTGTTGCTGAACCTTTTAAACCAAG	734		
DB 408	TTGAAGTGGCAACAGATATAAAGGAGCTTTAAATTTGTTGCTGAACCTTTTAAACCAAG	467		
QY 735	CAAAATCAATCTATTGTGCGAACAGATAGCGACAGAGGCGGAAACATTTGCTGTGCGA	794		
DB 468	CAAAATCAATCTATTGTGCGAACAGATAGCGACAGAGGCGGAAACATTTGCTGTGCGA	527		
QY 795	TCATTCATAAAGCAATGCTTTTCTAAGATATAAAGCTATATAAGCACTATCGATCAATA	854		
DB 528	TCATTCATAAAGCAATGCTTTTCTAAGATATAAAGCTATATAAGCACTATCGATCAATA	587		
QY 855	GTTTAGAAAAGATGTGATCCGTAGCGGTTTTCAAAATTTTGAACAGCAATGAAATTAAT	914		
DB 588	GTTTAGAAAAGATGTGATCCGTAGCGGTTTTCAAAATTTTGAACAGCAATGAAATTAAT	647		
QY 915	ATCCCTTTTATCAAGAGCGCAAAAGAAAACGAAATGATACACCAATCAGTGCAGAAA	974		
DB 648	ATCCCTTTTATCAAGAGCGCAAAAGAAAACGAAATGATACACCAATCAGTGCAGAAA	707		
QY 975	AGATATAATGGGAGATAAGACGGTTCGTGTTTCGTGCTGACCTTGCACCATATCATATAAT	1034		
DB 708	AGATATAATGGGAGATAAGACGGTTCGTGTTTCGTGCTGACCTTGCACCATATCATATAAT	767		
QY 1035	CGAAACAGCAAGAAATGGCGGAAAACGTAAAGAAATGATGGAATAAGACTTAGAAGCAA	1094		
DB 768	CGAAACAGCAAGAAATGGCGGAAAACGTAAAGAAATGATGGAATAAGACTTAGAAGCAA	827		
QY 1095	ACTTAAAGATGTGTCATAGTCAGATCTTAAATTTTGTATATAGGAAATTTGAAGTTA	1154		
DB 828	ACTTAAAGATGTGTCATAGTCAGATCTTAAATTTTGTATATAGGAAATTTGAAGTTA	887		
QY 1155	AATTAGATGTAAAAATTTGTAAATTAAGAGGAGTGAATACATGAAACAAAAATATAAAT	1214		
DB 888	AATTAGATGTAAAAATTTGTAAATTAAGAGGAGTGAATACATGAAACAAAAATATAAAT	947		
QY 1215	ATTCTCAAACTTTTAAACGGTGAAGAGTACTCAACCAATTAATAAACAATTTGAAT	1274		
DB 948	ATTCTCAAACTTTTAAACGGTGAAGAGTACTCAACCAATTAATAAACAATTTGAAT	1007		
QY 1275	TAAAGAAAACCGATACCGTTTACGAAATTTGGAACAGGTAAAGGGCATTTTAAACGAGAAC	1334		
DB 1008	TAAAGAAAACCGATACCGTTTACGAAATTTGGAACAGGTAAAGGGCATTTTAAACGAGAAC	1067		

QY 1335	TGGCTAAAATAAGTAAACAGGTAACGTCCTATTGAATTAGACAGATCATCTATTCAACTTAT	1394
DB 1068	TGGCTAAAATAAGTAAACAGGTAACGTCCTATTGAATTAGACAGATCATCTATTCAACTTAT	1127
QY 1395	CGTCAGAAAAATTAACAACTGATCTGCTGTCACCTTTAAATTCACCAAGATATTTACAGT	1454
DB 1128	CGTCAGAAAAATTAACAACTGATCTGCTGTCACCTTTAAATTCACCAAGATATTTACAGT	1187
QY 1455	TTCAATTCCTTAAACAAACAGAGGTATAAATTTGTTGGGAGTATTCCTTACCATTAAAGCA	1514
DB 1188	TTCAATTCCTTAAACAAACAGAGGTATAAATTTGTTGGGAGTATTCCTTACCATTAAAGCA	1247
QY 1515	CACAAATTTATAAAAAAGTGTGTTTTGAAAGCCATGCGTCTGACATCTATCTGATTTGTG	1574
DB 1248	CACAAATTTATAAAAAAGTGTGTTTTGAAAGCCATGCGTCTGACATCTATCTGATTTGTG	1307
QY 1575	AAGAGGATTTCTACAAAGGTACCTTTGATATTTACCGCACTAGGTTGCTCTTGACA	1634
DB 1308	AAGAGGATTTCTACAAAGGTACCTTTGATATTTACCGCACTAGGTTGCTCTTGACA	1367
QY 1635	CTCAAGTCTCGATTGAGCAATTTGCTTAAAGCTGCCAGCGGAATGCTTCACTTAAACCA	1694
DB 1368	CTCAAGTCTCGATTGAGCAATTTGCTTAAAGCTGCCAGCGGAATGCTTCACTTAAACCA	1427
QY 1695	AAGTAAACAGTGTCTTAATAAAGCTTACCCGCCATACACAGATGTTCCAGATAAATATT	1754
DB 1428	AAGTAAACAGTGTCTTAATAAAGCTTACCCGCCATACACAGATGTTCCAGATAAATATT	1487
QY 1755	GGAAAGCTATATACGCTACTTTGTTTCAAAATGGGTCAATCGAGATATCGTCAACTGTTTA	1814
DB 1488	GGAAAGCTATATACGCTACTTTGTTTCAAAATGGGTCAATCGAGATATCGTCAACTGTTTA	1547
QY 1815	CTAAAAATCAGTTTCTACAGCAATGAAAACAGCGCAAGTAAACAAATTTAAGTACCGTTA	1874
DB 1548	CTAAAAATCAGTTTCTACAGCAATGAAAACAGCGCAAGTAAACAAATTTAAGTACCGTTA	1607
QY 1875	CTTATGAGCAAGTATTTGCTATTTTAAATAGTATCTATTATTAAACGGGAGGAAATTAAT	1934
DB 1608	CTTATGAGCAAGTATTTGCTATTTTAAATAGTATCTATTATTAAACGGGAGGAAATTAAT	1667
QY 1935	TCTATGAGTCCGCTTTGTAATTTGGAAGTTACACGTTACTAAAGGGAATTTAGATATAA	1994
DB 1668	TCTATGAGTCCGCTTTGTAATTTGGAAGTTACACGTTACTAAAGGGAATTTAGATATAA	1727
QY 1995	TTATTAGTATACCTACTGACAGCTTCCAAAGGAGCTTAAGAGTCCCTAGCGCTCTTATCA	2054
DB 1728	TTATTAGTATACCTACTGACAGCTTCCAAAGGAGCTTAAGAGTCCCTAGCGCTCTTATCA	1787
QY 2055	TGGGGAAGCTCCGATCATATGCAAGACAAAATAAACTCGCAACAGCACTTTGGAGAAATGG	2114
DB 1788	TGGGGAAGCTCCGATCATATGCAAGACAAAATAAACTCGCAACAGCACTTTGGAGAAATGG	1847
QY 2115	GACGAATCGAGAAAACCTCTTTTACGCTGATTCATATCTAATAAGCCGTAAGGAGAC	2174
DB 1848	GACGAATCGAGAAAACCTCTTTTACGCTGATTCATATCTAATAAGCCGTAAGGAGAC	1907
QY 2175	GGGTTCAAAAAGGTTTAAATAAGGAGAGCAATCAATGATTTAGCTAGTAATATTTT	2234
DB 1908	GGGTTCAAAAAGGTTTAAATAAGGAGAGCAATCAATGATTTAGCTAGTAATATTTT	1967
QY 2235	TTGACAAACGTPGGAGAAATTTAGAGAACGTCCTTCCAAAGCCAGTTTCAAGAGCTAGTG	2294
DB 1968	TTGACAAACGTPGGAGAAATTTAGAGAACGTCCTTCCAAAGCCAGTTTCAAGAGCTAGTG	2027
QY 2295	CACATAACATAATTTTAAACGCTATAAGTGTGGAACACCTGTATATATATGAAAAAGCCG	2354
DB 2028	CACATAACATAATTTTAAACGCTATAAGTGTGGAACACCTGTATATATATGAAAAAGCCG	2087
QY 2355	TAGAAGAAATTAAGGAGGAGGAGAAATTTAGAGAAAGATTTTAAATGCCATATGCGGCGGT	2414
DB 2088	TAGAAGAAATTAAGGAGGAGGAGAAATTTAGAGAAAGATTTTAAATGCCATATGCGGCGGT	2147
QY 2415	TAGATGGGAACATATCAATTTTCTTGGAGATACAAATTTTGAAGGATTTACATGACACTG	2474







4635 AGATTTAATGCTTATATTAGCGAAGGAGGATGATATACAGCCCTATTAGTACGAC 4694  
4368 AGATTTAATGCTTATATTAGCGAAGGAGGATGATATACAGCCCTATTAGTACGAC 4427  
4695 CAAAAAGAGATTAGAGAAAGTGTAGGCATTCCTGAAACGACATAGATAAAATTTGCTGAA 4754  
4428 CAAAAAGAGATTAGAGAAAGTGTAGGCATTCCTGAAACGACATAGATAAAATTTGCTGAA 4487  
4755 GGTACTGAAGCGAATCAGAAATTTCTTTAGATTAAACAGGAGAAATGCTGGCAT 4814  
4488 GGTACTGAAGCGAATCAGAAATTTCTTTAGATTAAACAGGAGAAATGCTGGCAT 4547  
4815 TCAACTTGTAGTGTAAATTCATTTGCTGATCGATCAATTAAGTAAAGGAGAAAGGAA 4874  
4548 TCAACTTGTAGTGTAAATTCATTTGCTGATCGATCAATTAAGTAAAGGAGAAAGGAA 4607  
4875 AGAAAGCTATATAAGCGCTGACAAATTTCTTTGACTTAGAGATACATTCATCAAGA 4934  
4608 AGAAAGCTATATAAGCGCTGACAAATTTCTTTGACTTAGAGATACATTCATCAAGA 4667  
4935 GACTTTAAACAGCTAGCAAGCGCTTAAACGAGACACACTCGATTTGTTAGCTA 4994  
4668 GACTTTAAACAGCTAGCAAGCGCTTAAACGAGACACACTCGATTTGTTAGCTA 4727  
4995 TGATACAGGCTGAAATTAACCCGCACTATGCCATTACATTTATATCTATGATAGCTGT 5054  
4728 TGATACAGGCTGAAATTAACCCGCACTATGCCATTACATTTATATCTATGATAGCTGT 4787  
5055 TTGTTTTTTCTTCTGCTTTAGCGAATGATTAGCAGAAATATACAGATGAGATTTAAT 5114  
4788 TTGTTTTTTCTTCTGCTTTAGCGAATGATTAGCAGAAATATACAGATGAGATTTAAT 4847  
5115 TAATTTATAGGGGAGAGAGAGAGTAGCCGAAACCTTTTAGTTGGCTTGACCTGAAC 5174  
4848 TAATTTATAGGGGAGAGAGAGAGTAGCCGAAACCTTTTAGTTGGCTTGACCTGAAC 4907  
5175 GAAGTGAAGGAGAGAGAGAGAGTAGCCGAAACCTTTTAGTTGGCTTGACCTGAAC 5234  
4908 GAAGTGAAGGAGAGAGAGAGAGTAGCCGAAACCTTTTAGTTGGCTTGACCTGAAC 4967  
5235 TTTTAAATTTCTATCTTTTATAGGTCATTAGATGATATCTTTGCTCTATAAATTT 5294  
4968 TTTTAAATTTCTATCTTTTATAGGTCATTAGATGATATCTTTGCTCTATAAATTT 5027  
5295 TAGCAGCAATATAGATTTATGATAGGTCATTTAAGTTGAGCATATTAGAGAGAGAAA 5354  
5028 TAGCAGCAATATAGATTTATGATAGGTCATTTAAGTTGAGCATATTAGAGAGAGAAA 5087  
5355 TCTTGAAGATATTTGAAGACCGGATTCATGATGATGATGATGATGATGATGATGAT 5414  
5088 TCTTGAAGATATTTGAAGACCGGATTCATGATGATGATGATGATGATGATGATGAT 5147  
5415 GGTTTTAACTAAAGTAGTGAATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 5474  
5148 GGTTTTAACTAAAGTAGTGAATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 5207  
5475 TTGCTAGTCAAGTGATTAATA 5497  
5208 TTGCTAGTCAAGTGATTAATA 5230

RESULT 6  
AA15041  
ID AA15041 standard; DNA; 5230 BP.  
XX  
AC AA15041;  
XX  
XX  
DT 21-AUG-2000 (first entry)  
XX  
XX Nucleotide sequence of expression plasmid pTREX1.  
DE  
XX Intestinal mucosa; cytokine; cytokine antagonist; anti-inflammatory;  
KW

Gran-positive bacteria; Lactococcus lactis; gastrointestinal tract;  
inflammatory bowel disease; chronic colitis; Crohn's disease; IL10;  
ulcerative colitis; tumour necrosis factor; TNF; interleukin 10; ss.  
Synthetic.  
WO200023471-A2.  
27-APR-2000.  
06-OCT-1999; 99WO-EP007800.  
20-OCT-1998; 98EP-00203529.  
(VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.  
Steidler L, Remaut ER, Fiers W;  
WPI; 2000-422481/36.  
Treating inflammatory bowel diseases comprises administering a  
composition comprising a cytokine-producing or cytokine antagonist-  
producing Gram-positive bacterial strain.  
Disclosure; Fig 1B; 45pp; English.  
The present sequence represents an expression plasmid for use in the  
course of the invention. The specification describes an administration  
strategy for the delivery of the intestinal mucosa of cytokines or  
cytokine antagonists, preferably of acid sensitive anti-inflammatory  
agents such as interleukin 10 (IL10) or tumour necrosis factor (TNF). The  
method uses a cytokine-producing or cytokine antagonist-producing Gram-  
positive bacterial strain (such as Lactococcus lactis). The use of non-  
colonizing bacteria expressing cytokines or cytokine antagonists allows  
the treatment to be directed to the disease site, whilst minimizing the  
possibility of degradation along the gastrointestinal tract. The  
recombinant bacteria are used in the treatment of inflammatory bowel  
diseases, especially chronic colitis, Crohn's disease or an ulcerative  
colitis  
Sequence 5230 BP; 1930 A; 796 C; 982 G; 1522 T; 0 U; 0 Other;

Query Match 90.8%; Score 4991.8; DB 3; Length 5230;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 4996; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 495 TCTAATAGTAGTCCGCTGCTAACAAAGCCGAAAGAGAGCTGAGTTGGCTGCGCA 554  
DB 228 TCTCTAGATGATCCGCTGCTAACAAAGCCGAAAGAGAGCTGAGTTGGCTGCGCA 287  
QY 555 CGCTGAGCAATACTAGCATAAACCCCTTGGGGCTCTAAACGGGCTCTGAGGGGTTTTT 614  
DB 288 CGCTGAGCAATACTAGCATAAACCCCTTGGGGCTCTAAACGGGCTCTGAGGGGTTTTT 347  
QY 615 TGCTGAAAGGAGGAACTATATCCGGATGACCTGAGCAAGCTCTAGAAATCGATACGATT 674  
DB 348 TGCTGAAAGGAGGAACTATATCCGGATGACCTGAGCAAGCTCTAGAAATCGATACGATT 407  
QY 675 TTGAAGTGGCAACAGATAAAAAAGAGTTTAAATTTGTTGCTGAACTTTTAAACCAAG 734  
DB 408 TTGAAGTGGCAACAGATAAAAAAGAGTTTAAATTTGTTGCTCAACTTTTAAACCAAG 467  
QY 735 CAATACATCATTTGCTCCACAGATAGCGAGAGGCGGAAACATTCCTCGTGGTCA 794  
DB 468 CAATACATCATTTGCTCCACAGATAGCGAGAGGCGGAAACATTCCTCGTGGTCA 527  
QY 795 TCATTATAAAGCAATATCCCTTTTCTAAAGATAAAACGATATAAAGAGCTATGGATCAATA 854  
DB 528 TCATTATAAAGCAATATCCCTTTTCTAAAGATAAAACGATATAAAGAGCTATGGATCAATA 587  
QY 855 GTTTAGAAAGAGATGATCGTAGCGGTTTTTCAAAATTTGCAACCCAGAGATGATTAAT 914  
DB 588 GTTTAGAAAGAGATGATCGTAGCGGTTTTTCAAAATTTGCAACCCAGAGATGATTAAT 647

QY 915 ATCCCTTTTATCAAGAGCGCAAAAGAAAAAGAAATGATACACCAATCAGTGCAGAAAAA 974  
 Db 648 ATCCCTTTTATCAAGAGCGCAAAAGAAAAAGAAATGATACACCAATCAGTGCAGAAAAA 707  
 QY 975 AGATATAATGGAGAGATAGACGGTTTGGTTTGGTGTGCTGCACTTGCCACCATATCAAAAAAT 1034  
 Db 708 AGATATAATGGAGAGATAGACGGTTTGGTTTGGTGTGCTGCACTTGCCACCATATCAAAAAAT 767  
 QY 1035 CGAAACACCAAGAAATGGCGGAACGTAAGAAAGAGTTATGGAATAAGACTTAGAACCAA 1094  
 Db 768 CGAAACACCAAGAAATGGCGGAACGTAAGAAAGAGTTATGGAATAAGACTTAGAACCAA 827  
 QY 1095 ACTTAAGAGTGTGTGATGATGCAATCTTTAAATTTTGTATTAATAGGAATGAAGTTA 1154  
 Db 828 ACTTAAGAGTGTGTGATGATGCAATCTTTAAATTTTGTATTAATAGGAATTTGAAGTTA 887  
 QY 1155 AATTAGATGCTAAAAATTTGTAATTAAGAGGAGTGAATCATCAACAAATTAATAAATTAAT 1214  
 Db 888 AATTAGATGCTAAAAATTTGTAATTAAGAGGAGTGAATCATCAACAAATTAATAAATTAAT 947  
 QY 1215 ATTCTCAAAACCTTTTAAACGAGTGAAAGAGTACTCAACCAATTAATAAACAATTTGAATT 1274  
 Db 948 ATTCTCAAAACCTTTTAAACGAGTGAAAGAGTACTCAACCAATTAATAAACAATTTGAATT 1007  
 QY 1275 TAAAGAAACCGATACCGTTTACGAAATTTGAAACAGGTAAAGGCGATTTAAACGAGAAAC 1334  
 Db 1008 TAAAGAAACCGATACCGTTTACGAAATTTGAAACAGGTAAAGGCGATTTAAACGAGAAAC 1067  
 QY 1335 TGGCTAAATTAAGTAAACAGAGTAACTGCTATTTGAATTAAGAGTCACTCTATTTCACTTAT 1394  
 Db 1068 TGGCTAAATTAAGTAAACAGAGTAACTGCTATTTGAATTAAGAGTCACTCTATTTCACTTAT 1127  
 QY 1395 GGTCAAGAAATTAATACTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1454  
 Db 1128 GGTCAAGAAATTAATACTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1187  
 QY 1455 TTCAATTCCTTAAACAAACAGAGTAAATTTGTTGGAGTATTTCCATTTAAAGCA 1514  
 Db 1188 TTCAATTCCTTAAACAAACAGAGTAAATTTGTTGGAGTATTTCCATTTAAAGCA 1247  
 QY 1515 CACAAATTTATTAAGAGTGTGTTTTGAAAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1574  
 Db 1248 CACAAATTTATTAAGAGTGTGTTTTGAAAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1307  
 QY 1575 'AAGAGGATTTCAAGAGTACCTTGGATATTTCCGGAACACTAGGTTGCTCTTGCA 1634  
 Db 1308 AAGAGGATTTCAAGAGTACCTTGGATATTTCCGGAACACTAGGTTGCTCTTGCA 1367  
 QY 1635 CTCAGTCTCGATTCAGCAATTTGCTTTAGCTGCCAGCGGAATGCTTTATCTTAAACCAA 1694  
 Db 1368 CTCAGTCTCGATTCAGCAATTTGCTTTAGCTGCCAGCGGAATGCTTTATCTTAAACCAA 1427  
 QY 1695 AAGTAAACAGTGTCTTAATAAATTTACCCGCAATCCAGAGATTTCCAGATTAATTT 1754  
 Db 1428 AAGTAAACAGTGTCTTAATAAATTTACCCGCAATCCAGAGATTTCCAGATTAATTT 1487  
 QY 1755 GGAAGCTATATACGTAATTTGTTTCAAAATGGGTCAATTCGAGATATCGTCAACTGTTTA 1814  
 Db 1488 GGAAGCTATATACGTAATTTGTTTCAAAATGGGTCAATTCGAGATATCGTCAACTGTTTA 1547  
 QY 1815 CTAAAAATCAGTTTCATCAAGCAATGAAACAGCGCAAGTAAACAAATTTAAGTACCGTTTA 1874  
 Db 1548 CTAAAAATCAGTTTCATCAAGCAATGAAACAGCGCAAGTAAACAAATTTAAGTACCGTTTA 1607  
 QY 1875 CTTATGAGCAAGTATTTGTTTAAATAGTATTTTAAATTTTAAACCGGAGGAAATAAT 1934  
 Db 1608 CTTATGAGCAAGTATTTGTTTAAATAGTATTTTAAATTTTAAACCGGAGGAAATAAT 1667  
 QY 1935 TCTATGAGTCGCTTTGTTGAAATTTTGGAAAGTTTACAGTTTCTAAAGGGAATGTAGATAAA 1994  
 Db 1668 TCTATGAGTCGCTTTGTTGAAATTTTGGAAAGTTTACAGTTTCTAAAGGGAATGTAGATAAA 1727

QY 1995 TTATTAGGTATATCTACTGACAGCTTCCAAAGGAGCTTAAAGAGGTCCCTTAGCGCTCTTATCA 2054  
 Db 1728 TTATTAGGTATATCTACTGACAGCTTCCAAAGGAGCTTAAAGAGGTCCCTTAGCGCTCTTATCA 1787  
 QY 2055 TGGGGAAGCTCCGATCATATGCAAGCAAAATAAACTCGCAACAGCACTTCGAGAAATGG 2114  
 Db 1788 TGGGGAAGCTCCGATCATATGCAAGCAAAATAAACTCGCAACAGCACTTCGAGAAATGG 1847  
 QY 2115 GACGAATCGAGAAAAACCTCTTTTACGCTGGATTTACATATCTTAATAAAGCCGTAGAGAC 2174  
 Db 1848 GACGAATCGAGAAAAACCTCTTTTACGCTGGATTTACATATCTTAATAAAGCCGTAGAGAC 1907  
 QY 2175 GGGTTCAAAAAGGTTTAAATAAGGAGAGCAATCAATGCAATTTAGCTAGACTATATTTT 2234  
 Db 1908 GGGTTCAAAAAGGTTTAAATAAGGAGAGCAATCAATGCAATTTAGCTAGACTATATTTT 1967  
 QY 2235 TTGGACAAACGCTGGAGAAATTTAGAGAACGCTCTCTCAAGACCCAGTTTCAAAAGAGCTAGTG 2294  
 Db 1968 TTGGACAAACGCTGGAGAAATTTAGAGAACGCTCTCTCAAGACCCAGTTTCAAAAGAGCTAGTG 2027  
 QY 2295 CACTAAACATTAATTAACGCTTAAGTGTGTGGAACACTGTATATATGAAAGAGCG 2354  
 Db 2028 CACTAAACATTAATTAACGCTTAAGTGTGTGGAACACTGTATATATGAAAGAGCG 2087  
 QY 2355 TAGAAGAAATTAAGCAAGAGAGAAATTTAGAGAAATTTAATGCCATATGCTGGCCGT 2414  
 Db 2088 TAGAAGAAATTAAGCAAGAGAGAAATTTAGAGAAATTTAATGCCATATGCTGGCCGT 2147  
 QY 2415 TAGGATGGAAACATATCAATTTTCTTGAGAAATACAAATTTTGAGGATTAATGACACTG 2474  
 Db 2148 TAGGATGGAAACATATCAATTTTCTTGAGAAATACAAATTTTGAGGATTAATGACACTG 2207  
 QY 2475 GGCAAATGAATTTAGCTCTTTTACGTATAAAGAGCCGTTTTATTCTTAATTAACGGCT 2534  
 Db 2208 GGCAAATGAATTTAGCTCTTTTACGTATAAAGAGCCGTTTTATTCTTAATTAACGGCT 2267  
 QY 2535 CTTTATTAGAAAAAATCCTTAGCGTGTGTTTTTTTCGAAATGCTGCGGTACCCCAAGA 2594  
 Db 2268 CTTTATTAGAAAAAATCCTTAGCGTGTGTTTTTTTCGAAATGCTGCGGTACCCCAAGA 2327  
 QY 2595 ATTAGAATGAGTATGATCAATTTTACGATAGAAATCAGAAATCAGATCCAAACCAT 2654  
 Db 2328 ATTAGAATGAGTATGATCAATTTTACGATAGAAATCAGAAATCAGATCCAAACCAT 2387  
 QY 2655 AAAAAACCTAGAACAAATTTGCAAGTTTAACTCAACGCTAGTGGATTTTAAATCCC 2714  
 Db 2388 AAAAAACCTAGAACAAATTTGCAAGTTTAACTCAACGCTAGTGGATTTTAAATCCC 2447  
 QY 2715 AAATGAGCCACAGAACCGAGCCAGAACAGAAATCAGAAACAGTAAATTTGGATTTAGA 2774  
 Db 2448 AAATGAGCCACAGAACCGAGCCAGAACAGAAATCAGAAACAGTAAATTTGGATTTAGA 2507  
 QY 2775 AATGGAAGAAAGAAAAAGCAATGCTTGTGTAATTAATGCAAGAAATCGTTGCTTATTT 2834  
 Db 2508 AATGGAAGAAAGAAAAAGCAATGCTTGTGTAATTAATGCAAGAAATCGTTGCTTATTT 2567  
 QY 2835 TTTTTPAAAAGCGGTATAGATATAACGAAACAAACGAACTGAAATAGAAACGAAAAAG 2894  
 Db 2568 TTTTTPAAAAGCGGTATAGATATAACGAAACAAACGAACTGAAATAGAAACGAAAAAG 2627  
 QY 2895 AGCCATGACACATTTTATAAATGTTTGACGACATTTTAAATGATAGCCGATAGAT 2954  
 Db 2628 AGCCATGACACATTTTATAAATGTTTGACGACATTTTAAATGATAGCCGATAGAT 2687  
 QY 2955 TGCCAAACCAACGCTTATCAGTTAGTGAATGAACTCTTCCCTCGTATAGAGTTATTTAA 3014  
 Db 2688 TGCCAAACCAACGCTTATCAGTTAGTGAATGAACTCTTCCCTCGTATAGAGTTATTTAA 2747  
 QY 3015 TTAATCTTTGTTGAGAGCGGTATTAACCGTACTACTATATATAGGGAATCAGAGAGT 3074  
 Db 2748 TTAATCTTTGTTGAGAGCGGTATTAACCGTACTACTATATATAGGGAATCAGAGAGT 2807  
 QY 3075 TTTCAAGTATCTAAGCTACTGAATTTAAGAAATTTGTAAGCAATCAATCGGAAATCGTTTG 3134

Db 2808 TTTCAGATCTAAGCTACTGAATTTAAGAAATTTGTTAAGCAATCAATCGGAATCGTTTG 2867  
Qy 3135 ATTGCTTTTGTGATTCATTTATAGAAAGTGGAGTTGTGATGATCAATGATGAATGTA 3194  
Db 2868 ATTGCTTTTGTGATTCATTTATAGAAAGTGGAGTTGTGATGATCAATGATGAATGTA 2927  
Qy 3195 AACCTATATAAAATAGTTTATTTGGAGATAAGAAAATAGCAAAATATCTATACACTAGA 3254  
Db 2928 AACCTATATAAAATAGTTTATTTGGAGATAGAAAATAGCAAAATATCTATACACTAGA 2987  
Qy 3255 AACCTTTAAGAAAGTTAGAAAGAGAAATATCTACTTAGAAACAAAATCAGATAGTA 3314  
Db 2988 AACCTTTAAGAAAGTTAGAAAGAGAAATATCTACTTAGAAACAAAATCAGATAGTA 3047  
Qy 3315 TTTTCTTCGAGGGGGAAGTATATATATAGTTAATAGAAAATAACAAAATATATTA 3374  
Db 3048 TTTTCTTCGAGGGGGAAGTATATATATAGTTAATAGAAAATAACAAAATATATTA 3107  
Qy 3375 TTCGATTTAGTGGAAAAATTTGACTTATAAGGAGAAAAATCTTTTCAAAACATGCAAT 3434  
Db 3108 TTCGATTTAGTGGAAAAATTTGACTTATAAGGAGAAAAATCTTTTCAAAACATGCAAT 3167  
Qy 3435 ATTGAACAGTTGNAATGAAAAGCAACCAAGTTAATTAACCACTTATTTATAGATT 3494  
Db 3168 ATTGAACAGTTGAATGAAAAGCAACCAAGTTAATTAACCACTTATTTATAGATT 3227  
Qy 3495 TATAGGAAAGAGAACAGCTGAATGAATATCCCTTTTGTAGAACTGTGCTTCATGA 3554  
Db 3228 TATAGGAAAGAGAACAGCTGAATGAATATCCCTTTTGTAGAACTGTGCTTCATGA 3287  
Qy 3555 CGGCTTGTTAAAGTACAAATTTAAATATAGTAAATTCGCTCAATCACTACCAAGCCAGG 3614  
Db 3288 CGGCTTGTTAAAGTACAAATTTAAATATAGTAAATTCGCTCAATCACTACCAAGCCAGG 3347  
Qy 3615 TAAAGCAAGGGGCTATTTTGGTATCGCTCAAAATCAAGCATGATTCGCGGTGCTGG 3674  
Db 3348 TAAAGCAAGGGGCTATTTTGGTATCGCTCAAAATCAAGCATGATTCGCGGTGCTGG 3407  
Qy 3675 TGTGTTCTGACTCCGAGGAAGCGATTCAAGAAAATCAAGATACATTTACACATTGGAC 3734  
Db 3408 TGTGTTCTGACTCCGAGGAAGCGATTCAAGAAAATCAAGATACATTTACACATTGGAC 3467  
Qy 3735 ACCCAAGCTTATCGTTATGMAAGTATGAGCGAAACCGTTTCATACAGAAAGACA 3794  
Db 3468 ACCCAAGCTTATCGTTATGMAAGTATGAGCGAAACCGTTTCATACAGAAAGACA 3527  
Qy 3795 TTCTGAAAACAATTTAAGACAAATCAATACCTTCTTTATTTGATTTTCAACAGGC 3854  
Db 3528 TTCTGAAAACAATTTAAGACAAATCAATACCTTCTTTATTTGATTTTCAACAGGC 3587  
Qy 3855 AAAAGAACTATTTACAGAGCGATATTTTAAACCGCTATTTGATTTAGTTTATGCC 3914  
Db 3588 AAAAGAACTATTTACAGAGCGATATTTTAAACCGCTATTTGATTTAGTTTATGCC 3647  
Qy 3915 TACTATGATTTCAAAATCTGATAAGGTATCAAGCATATTTTGTTTTGAACCGCCAGT 3974  
Db 3648 TACTATGATTTCAAAATCTGATAAGGTATCAAGCATATTTTGTTTTGAACCGCCAGT 3707  
Qy 3975 CTATGTGACTTCAAAATCAGAAATTTAAATCTGTCAAGCGACCAAAATATTTTCGAAAA 4034  
Db 3708 CTATGTGACTTCAAAATCAGAAATTTAAATCTGTCAAGCGACCAAAATATTTTCGAAAA 3767  
Qy 4035 TATCGGAGATATTTTCGAAAGTCTTTGCCAGTTGATCTAACGTTAATCATTTTGGTAT 4094  
Db 3768 TATCGGAGATATTTTCGAAAGTCTTTGCCAGTTGATCTAACGTTAATCATTTTGGTAT 3827  
Qy 4095 TGCTCGCATACCAAGAACGGACATGTAGAAATTTTTTGATCTCTAATTAACGTTATCTTT 4154  
Db 3828 TGCTCGCATACCAAGAACGGACATGTAGAAATTTTTTGATCTCTAATTAACGTTATCTTT 3887  
Qy 4155 CAAAGATGGCAAGTTGGCTTTTCAACACACAGATAAAGGCTTTACTCGTTCAAG 4214

Db 3888 CAAAGATGGCAAGATTTGGTCTTTTCAAAACACAGATAAAGGGCTTTTACTCGTTCAAG 3947  
Qy 4215 TCTAACGGTTTTTAAGCGGTACAGAGGCAAAACAAAGTAGATGAACCTGGTTTAAATCT 4274  
Db 3948 TCTAACGGTTTTTAAGCGGTACAGAGGCAAAACAAAGTAGATGAACCTGGTTTAAATCT 4007  
Qy 4275 CTTATTTGACGAAAAACGAAATTTTTCAGGAGAAAAAGGGTTTAAATAGGCGCTAATAACGTCAT 4334  
Db 4008 CTTATTTGACGAAAAACGAAATTTTTCAGGAGAAAAAGGGTTTAAATAGGCGCTAATAACGTCAT 4067  
Qy 4335 GTTTACCTCTCTTTAGCCTACTTTTAGTTTCAGGCTTATTCATTCGAAACGTCGGATATAA 4394  
Db 4068 GTTTACCTCTCTTTAGCCTACTTTTAGTTTCAGGCTTATTCATTCGAAACGTCGGATATAA 4127  
Qy 4395 TATGTTTCAGTTTAAATTCGATTAGATCAACCCCTTAGAGAAAAAGAAAGTAATCAAAAT 4454  
Db 4128 TATGTTTCAGTTTAAATTCGATTAGATCAACCCCTTAGAGAAAAAGAAAGTAATCAAAAT 4187  
Qy 4455 TGTTAGAGTGCCTTATTCAGAAAACTATCAAGGGCTAATAGGGAATACATTTACCATTTCT 4514  
Db 4188 TGTTAGAGTGCCTTATTCAGAAAACTATCAAGGGCTAATAGGGAATACATTTACCATTTCT 4247  
Qy 4515 TTGCAAGCTTCGGGTATCAAGTGATTTAAACAGTAAGATTTATTTGTCGCTCAAGGGTG 4574  
Db 4248 TTGCAAGCTTCGGGTATCAAGTGATTTAAACAGTAAGATTTATTTGTCGCTCAAGGGTG 4307  
Qy 4575 GTTTAAATTCAGAAAAAAGAACGTAACGTCCTCAACGTCCTCATTTGTCAGAAATGAAAGA 4634  
Db 4308 GTTTAAATTCAGAAAAAAGAACGTAACGTCCTCAACGTCCTCATTTGTCAGAAATGAAAGA 4367  
Qy 4635 AGATTTAATGCTTATATTTAGGAAAAAAGCGATGATACAGCCCTTATTTAGTCACGAC 4694  
Db 4368 AGATTTAATGCTTATATTTAGGAAAAAAGCGATGATACAGCCCTTATTTAGTCACGAC 4427  
Qy 4695 CAAAAAGAGATTAGAGAAAGTGTAGGCAATTCCTGAACGGACATTTAGATAAATTTCTCGAA 4754  
Db 4428 CAAAAAGAGATTAGAGAAAGTGTAGGCAATTCCTGAACGGACATTTAGATAAATTTCTCGAA 4487  
Qy 4755 GGTACTGAAGCGCAATCAGAAAATTTCTTTAAGATTTAAACAGGAAGAAATTCGTTGGCAT 4814  
Db 4488 GGTACTGAAGCGCAATCAGAAAATTTCTTTAAGATTTAAACAGGAAGAAATTCGTTGGCAT 4547  
Qy 4815 TCAACTGTCTAGTGTAAATCAATCTGCTATCGATCAATTAAGTAAAAAAGAAAGAAAAA 4874  
Db 4548 TCAACTGTCTAGTGTAAATCAATCTGCTATCGATCAATTAAGTAAAAAAGAAAGAAAAA 4607  
Qy 4875 AGAAAGCTATATAAAGGGCTGACAAATTCCTTTGACCTTAGAGCAATACATTCATTCAAGA 4934  
Db 4608 AGAAAGCTATATAAAGGGCTGACAAATTCCTTTGACCTTAGAGCAATACATTCATTCAAGA 4667  
Qy 4935 GACTTTAAACAGCTAGCAGACGCCCTTAAACGAGACACAACTCGATTTGTTTAGCTA 4994  
Db 4668 GACTTTAAACAGCTAGCAGACGCCCTTAAACGAGACACAACTCGATTTGTTTAGCTA 4727  
Qy 4995 TGATACAGGCTGAAAAATAAAAACCCGCACTATGCCATTTACATTTATCTATGATACGTTG 5054  
Db 4728 TGATACAGGCTGAAAAATAAAAACCCGCACTATGCCATTTACATTTATCTATGATACGTTG 4787  
Qy 5055 TTGTTTTTTTCTTTGCTTTTAGCGAATGATAGCAGAAATATACAGAGTAAGATTTTAAAT 5114  
Db 4788 TTCTTTTCTTTTCTTTGCTTTTAGCGAATGATTAGCAGAAATATACAGAGTAAGATTTTAAAT 4847  
Qy 5115 TAATTTATTTAGGGGAGAGAGAGAGTAGCCCGAAAAAATTTTAGTTGGCTTGGACTGAAC 5174  
Db 4848 TAATTTATTTAGGGGAGAGAGAGAGTAGCCCGAAAAAATTTTAGTTGGCTTGGACTGAAC 4907  
Qy 5175 GAAGTAGGGAAGGCTACTAAAAACGTCGAGGGCAGTGAGAGCGAAGCGAACACTTGAT 5234  
Db 4908 GAAGTAGGGAAGGCTACTAAAAACGTCGAGGGCAGTGAGAGCGAAGCGAACACTTGAT 4967  
Qy 5235 TTTTAAATTTTCTATCTTTTATAGTCAATAGAGATATCTTTATTTGCTCTATAAACTATT 5294  
Db 4968 TTTTAAATTTTCTATCTTTTATAGTCAATAGAGATATCTTTATTTGCTCTATAAACTATT 5027



Db 1293 GAAGAGGATTTACAGCGTACCTTGGATATTCACCGAACCTAGGGTTGCTCTTGAC 1352  
Qy 1634 ACTCAAGTCTCGATTACAGCAATGCTTAAGCTGCCAGCGGAATGCTTTTCATCTCAACCA 1693  
Db 1353 ACTCAAGTCTCGATTACAGCAATGCTTAAGCTGCCAGCGGAATGCTTTTCATCTCAACCA 1412  
Qy 1694 AAAGTAAACAGTGTCTTAATAAATCTTACCCGCCATACCAAGATGTTCCAGATAAATAT 1753  
Db 1413 AAAGTAAACAGTGTCTTAATAAATCTTACCCGCCATACCAAGATGTTCCAGATAAATAT 1472  
Qy 1754 TGAAGCTATATAGTACTGTTGTTTCAAAATGGTCAATCGAAGATATCGTCAACTGTTT 1813  
Db 1473 TGAAGCTATATAGTACTGTTGTTTCAAAATGGTCAATCGAAGATATCGTCAACTGTTT 1532  
Qy 1814 ACTAAAAATCAGTTTCACTCAAGCAATGAACACGCCCAAGATGAACAAATTTAAGTACCGTT 1873  
Db 1533 ACTAAAAATCAGTTTCACTCAAGCAATGAACACGCCCAAGATGAACAAATTTAAGTACCGTT 1592  
Qy 1874 ACTTATGAGCAAGTATGCTCTATTTTAAATAGTTATCTATATTTAACCGGAGGAATAA 1933  
Db 1593 ACTTATGAGCAAGTATGCTCTATTTTAAATAGTTATCTATATTTAACCGGAGGAATAA 1652  
Qy 1934 TTCTATGAGTCCGTTTGTAAATTTGGAAGTTTACAGCTTCTAAAGGGAATCTAGATAA 1993  
Db 1653 TTCTATGAGTCCGTTTGTAAATTTGGAAGTTTACAGCTTCTAAAGGGAATCTAGATAA 1712  
Qy 1994 ATTATAGGTATACCTGACAGCTTCCAAAGAGCTAAAGAGTCCCTAGCGCTCTTATC 2053  
Db 1713 ATTATAGGTATACCTGACAGCTTCCAAAGAGCTAAAGAGTCCCTAGCGCTCTTATC 1772  
Qy 2054 ATGGGAGCTCGGATCATATGCAAGCAAAATTAACCTGCAACAGCACTTGGAGAAATG 2113  
Db 1773 ATGGGAGCTCGGATCATATGCAAGCAAAATTAACCTGCAACAGCACTTGGAGAAATG 1832  
Qy 2114 GGAAGATCGAGAAACCCCTCTTACGCTGGATTACATATCTAATAAGCCCTGAAGGAGA 2173  
Db 1833 GGACGAATCGAGAAACCCCTCTTACGCTGGATTACATATCTAATAAGCCCTGAAGGAGA 1892  
Qy 2174 CGGGTTCAAAAGGTTTAAATAAGGAGAGCAATCAATGCATTAGCTAGACTATATTT 2233  
Db 1893 CGGGTTCAAAAGGTTTAAATAAGGAGAGCAATCAATGCATTAGCTAGACTATATTT 1952  
Qy 2234 TTTGGACAACCTGGAGAAATTTAGAGAAACGTCCTCCAGAGCAGTTACAAAGAGCTAGT 2293  
Db 1953 TTTGGACAACCTGGAGAAATTTAGAGAAACGTCCTCCAGAGCAGTTACAAAGAGCTAGT 2012  
Qy 2294 GCCTAAACATATATTAACGCTATAAGTGTGTGAAACACTGTATATATGGAAGAACCC 2353  
Db 2013 GCCTAAACATATATTAACGCTATAAGTGTGTGAAACACTGTATATATGGAAGAACCC 2072  
Qy 2354 GTAGAGAAATTAAGCAAGAGAGAAATTTAGAGAAATTTAATGCTATATGCTGGCCG 2413  
Db 2073 GTAGAGAAATTAAGCAAGAGAGAAATTTAGAGAAATTTAATGCTATATGCTGGCCG 2132  
Qy 2414 TTAGGATGGGAACATATCAATTTCTTGGAGAAATCAAAATTTGAAGGATTACATGACCT 2473  
Db 2133 TTAGGATGGGAACATATCAATTTCTTGGAGAAATCAAAATTTGAAGGATTACATGACCT 2192  
Qy 2474 GGGCAATGAAATTTACGCTTTTACGTTAAGAGAGCGGTTTATTTCTTAATATACGCG 2533  
Db 2193 GGGCAATGAAATTTACGCTTTTACGTTAAGAGAGCGGTTTATTTCTTAATATACGCG 2252  
Qy 2534 TCTTTTATAGAAAAATCTTACGCTGGTTTCTTCCGAATGCTGGCGGTACCCCAAG 2593  
Db 2253 TCTTTTATAGAAAAATCTTACGCTGGTTTCTTCCGAATGCTGGCGGTACCCCAAG 2312  
Qy 2594 AATTAGAAATGAGTAGATCAAAATTTTACAGAAATAGAAATCAGAAATCAGATCCAAACCA 2653  
Db 2313 AATTAGAAATGAGTAGATCAAAATTTTACAGAAATAGAAATCAGAAATCAGATCCAAACCA 2372  
Qy 2654 TAAAAACACTAGAAACAAATTTGCAAGGTTTAACTAACTCAACGCTAGTAGTGGATTTAATCC 2713

Db 2373 TAAAAACACTAGAAACAAATTTGCAAAATTTAACTCAACGCTAGTAGTGGATTTAAATCC 2432  
Qy 2714 CAAATGAGCCACAGAACCCAGAGCCAGAAACAGAAATCAGAAATCAGAAATCAGAAATTTAG 2773  
Db 2433 CAAATGAGCCACAGAACCCAGAGCCAGAAACAGAAATCAGAAATCAGAAATTTAG 2492  
Qy 2774 AAATGGAAGAAAGAAAGAAAGCAATCTTCGTTGTAATTAATGACGAAATCGTTGCTTAT 2833  
Db 2493 AAATGGAAGAAAGAAAGAAAGCAATCTTCGTTGTAATTAATGACGAAATCGTTGCTTAT 2552  
Qy 2834 TTTTTTTTAAAGCGGTACTAGATATAACGAACACAGCAATCTTAAATGCAATGAAACGAAATA 2893  
Db 2553 TTTTTTTTAAAGCGGTACTAGATATAACGAACACAGCAATCTTAAATGCAATGAAACGAAATA 2612  
Qy 2894 GAGCCATGACACATTTTAAATGTTTGCAGCAGCACTTTTAAATGCAATGCAATGCAATGCAAT 2953  
Db 2613 GAGCCATGACACATTTTAAATGTTTGCAGCAGCACTTTTAAATGCAATGCAATGCAATGCAAT 2672  
Qy 2954 TTGCCAAACCAACGCTTATCAGTTAGTGCAGATGAACCTTCCCTCGTGAAGAAATTTATTA 3013  
Db 2673 TTGCCAAACCAACGCTTATCAGTTAGTGCAGATGAACCTTCCCTCGTGAAGAAATTTATTA 2732  
Qy 3014 ATTAATCTTTGTTGAAGAGCGGTATATAACCGTACTTATCATTTATAGGGAATCAGAGAG 3073  
Db 2733 ATTAATCTTTGTTGAAGAGCGGTATATAACCGTACTTATCATTTATAGGGAATCAGAGAG 2792  
Qy 3074 TTTTCAAGTATCTAAGCTACTGAATTTAAGAAATTTAAGAAATTTAAGAAATTTAAGAAATTTA 3133  
Db 2793 TTTTCAAGTATCTAAGCTACTGAATTTAAGAAATTTAAGAAATTTAAGAAATTTAAGAAATTT 2852  
Qy 3134 GATTGCTTTTTTGTGTTTCAATTTATAGAGGTGGAGTTTGTATGAATCATCATCAATGTA 3193  
Db 2853 GATTGCTTTTTTGTGTTTCAATTTATAGAGGTGGAGTTTGTATGAATCATCATCAATGTA 2912  
Qy 3194 AAACCTTATATAAAATAGTTTATTTGGAGATAAGAAATTTAGCAAAATCTATACACTAG 3253  
Db 2913 AAACCTTATATAAAATAGTTTATTTGGAGATAAGAAATTTAGCAAAATCTATACACTAG 2972  
Qy 3254 AAACCTTTAAGAGAGTTAGAAAGAGAAATTTCTACTTTAGAAACAAATCAGATAAGT 3313  
Db 2973 AAACCTTTAAGAGAGTTAGAAAGAGAAATTTCTACTTTAGAAACAAATCAGATAAGT 3032  
Qy 3314 ATTTTCTTCCGAGGGGGAAGATTATATATATAAGTTTAAATAGAAATTAACAAATTAATTT 3373  
Db 3033 ATTTTCTTCCGAGGGGGAAGATTATATATATAAGTTTAAATAGAAATTAACAAATTAATTT 3092  
Qy 3374 ATTCGATTAGTGGAAAAATTTGACTTATAAAGGAAAAATCTTTTCAAAAATGCAATG 3433  
Db 3093 ATTCGATTAGTGGAAAAATTTGACTTATAAAGGAAAAATCTTTTCAAAAATGCAATG 3152  
Qy 3434 TATTGAAACAGTTGAATGAAAAAGCAACCAAGTTAAATTAACCAACCTTTTATAGAT 3493  
Db 3153 TATTGAAACAGTTGAATGAAAAAGCAACCAAGTTAAATTAACCAACCTTTTATAGAT 3212  
Qy 3494 TTATAGAAAGGAGAACAGCTGAATGAATATCCCTTTTGTGTAGAAATCTGTGCTCATG 3553  
Db 3213 TTATAGAAAGGAGAACAGCTGAATGAATATCCCTTTTGTGTAGAAATCTGTGCTCATG 3272  
Qy 3554 ACGGCTGTTTAAAGTACAAATTTAAATTAAGTAAATTCGCTCAATCACTACCAAGCAG 3613  
Db 3273 ACGGCTGTTTAAAGTACAAATTTAAATTAAGTAAATTCGCTCAATCACTACCAAGCAG 3332  
Qy 3614 GTAAAAAGCAAGGGCTATTTTTCGCTATCGCTCAAAATCAAGCATGATTGGCGGTGCTG 3673  
Db 3333 GTAAAAAGCAAGGGCTATTTTTCGCTATCGCTCAAAATCAAGCATGATTGGCGGTGCTG 3392  
Qy 3674 GTGTTGTTCTGACITTCGAGGAGCGATTCAAGAAATTCAGATAATCAATTTACATTTGGA 3733  
Db 3393 GTGTTGTTCTGACITTCGAGGAGCGATTCAAGAAATTCAGATAATCAATTTACATTTGGA 3452  
Qy 3734 CACCCAAAGTTTATCGTTTATGGAACGTTATGCAAGGAAACCGTTTCATACAGAAAGGAC 3793  
Db 3453 CACCCAAAGTTTATCGTTTATGGAACGTTATGCAAGGAAACCGTTTCATACAGAAAGGAC 3512



QY	3794	ATTCTGAAACAAATTAAGCAAAATCAATACCTCTCTTTATTTGATTTTGTATTTATTCACACGG	3853	QY	4874	AAGAAAGCTATATAAAGCGCTGACAAATCTTTTGACTTTAGACATACATTCATTCAAG	4933
DB	3513	ATTCTGAAACAAATTAAGCAAAATCAATACCTCTCTTTATTTGATTTTGTATTTATTCACACGG	3572	DB	4593	AAGAAAGCTATATAAAGCGCTGACAAATCTTTTGACTTTAGACATACATTCATTCAAG	4652
QY	3854	CAAAAGAACTATTTTCAGCAAGCGATATTTTAAACACCGCTATTGATTTAGGTTTATGC	3913	QY	4934	AGACTTTTAAACAGCTAGCAGAACCGCCCTAAACCGGACACCAACTCGATTGTTTGTAGCT	4993
DB	3573	CAAAAGAACTATTTTCAGCAAGCGATATTTTAAACACCGCTATTGATTTAGGTTTATGC	3632	DB	4653	AGACTTTTAAACAGCTAGCAGAACCGCCCTAAACCGGACACCAACTCGATTGTTTGTAGCT	4712
QY	3914	CTACTATGATTAATCTGATTAAGAGTTATCAAGCATATTTTGTGTTTATAGAAACGCCAG	3973	QY	4994	ATGATACAGCTGAAATATAAACCCGCACTATGCCATTACATTTTATATCTATGATACGTG	5053
DB	3633	CTACTATGATTAATCTGATTAAGAGTTATCAAGCATATTTTGTGTTTATAGAAACGCCAG	3692	DB	4713	ATGATACAGCTGAAATATAAACCCGCACTATGCCATTACATTTTATATCTATGATACGTG	4772
QY	3974	TCTATGCACTTCAAAATCAGAAATTAATCTGTCAAAGCAGCCAAATTAATTTTCGCAAA	4033	QY	5054	TTTGTGTTTCTTCTGCTGTTTACGCAATGATTAGCAGAAATATACAGAGTAAAGATTTTAA	5113
DB	3693	TCTATGCACTTCAAAATCAGAAATTAATCTGTCAAAGCAGCCAAATTAATTTTCGCAAA	3752	DB	4773	TTTGTGTTTCTTCTGCTGTTTACGCAATGATTAGCAGAAATATACAGAGTAAAGATTTTAA	4832
QY	4034	ATATCCGAGAATATTTTCGAAAGCTTTTGCAGTTGATCTAAACCGTGAATCAATTTTGGTA	4093	QY	5114	TTAATTTATGGGGGAGAGGAGAGTAGCCCGAAAACTTTTAGTTGGCTTGGACTGAA	5173
DB	3753	ATATCCGAGAATATTTTGCAGAGCTTTTGCAGTTGATCTAAACCGTGAATCAATTTTGGTA	3812	DB	4833	TTAATTTATGGGGGAGAGGAGAGTAGCCCGAAAACTTTTAGTTGGCTTGGACTGAA	4892
QY	4094	TTGCTCGCATACCAAGACGACATGTAGAAATTTTGTATCTTAATTTACGTTTATCTT	4153	QY	5174	CGAAGTGAAGGAAAGCGCTACTATAAAACGTCGAGGGGCACTGAGAGCGAAACCACTTGA	5233
DB	3813	TTGCTCGCATACCAAGACGACATGTAGAAATTTTGTATCTTAATTTACGTTTATCTT	3872	DB	4893	CGAAGTGAAGGAAAGCGCTACTATAAAACGTCGAGGGGCACTGAGAGCGAAACCACTTGA	4952
QY	4154	TCAAAGAAATGCAAGATTTGCTTTCAAACCAACAGATAATTAAGGGCTTTTACTCGTTCAA	4213	QY	5234	TTTGTGTTTCTTCTATCTTTTATAGTCTATAGTACTATTTTGTCTTATAAATCTAT	5293
DB	3873	TCAAAGAAATGCAAGATTTGCTTTCAAACCAACAGATAATTAAGGGCTTTTACTCGTTCAA	3932	DB	4953	TTTGTGTTTCTTCTATCTTTTATAGTCTATAGTACTATTTTGTCTTATAAATCTAT	5012
QY	4214	GTCATACGGTTTAAAGCGGTACAGAACCAAAACAGTAGATGAACCTGTTTAAATC	4273	QY	5294	TTAGCAGCATATAGATTTTATGAATAGGTCATTTAAGTTGAGCATATTAGAGGAGGAAA	5353
DB	3933	GTCATACGGTTTAAAGCGGTACAGAACCAAAACAGTAGATGAACCTGTTTAAATC	3992	DB	5013	TTAGCAGCATATAGATTTTATGAATAGGTCATTTAAGTTGAGCATATTAGAGGAGGAAA	5072
QY	4274	TCTTATTCGCAAGAAATTTTTCAGGAGAAAGGGTTTAAAGGGCGTAAATACGTC	4333	QY	5354	ATCTTGGAGAAATTTTGAAGAACCGGATTACATGGATTGATTGTTGTTGTTAGTA	5413
DB	3993	TCTTATTCGCAAGAAATTTTTCAGGAGAAAGGGTTTAAAGGGCGTAAATACGTC	4052	DB	5073	ATCTTGGAGAAATTTTGAAGAACCGGATTACATGGATTGATTGTTGTTGTTAGTA	5132
QY	4334	TGTTTACCGTCTCTTTAGCTTCTTTAGTTCAGGCTATTCATCGAAACGTCGCAATATA	4393	QY	5414	TGCTTTTAACTAAAAGTAGTGAATTTTGTGTTGTTGTTGTTGTTGTTAGTA	5473
DB	4053	TGTTTACCGTCTCTTTAGCTTCTTTAGTTCAGGCTATTCATCGAAACGTCGCAATATA	4112	DB	5133	TGCTTTTAACTAAAAGTAGTGAATTTTGTGTTGTTGTTGTTGTTGTTAGTA	5192
QY	4394	ATATGTTTGTAGTTAATTAATCGATTAGATCAACCTTAGAAGAAAGAGTAAATCAAAA	4453	QY	5474	TTTGCTAGTCAAAAGTGAATTAATA	5497
DB	4113	ATATGTTTGTAGTTAATTAATCGATTAGATCAACCTTAGAAGAAAGAGTAAATCAAAA	4172	DB	5193	TTTGCTAGTCAAAAGTGAATTAATA	5216
QY	4454	TTGTTAGAGTGCCTATTCAAGAAACCTATCAAGGGCTTAATAGGAAATACATTAACATTC	4513	RESULT 8			
DB	4173	TTGTTAGAGTGCCTATTCAAGAAACCTATCAAGGGCTTAATAGGAAATACATTAACATTC	4232	AX07550			
QY	4514	TTTGCAAGCTGGGTATCAGTGAATTTAAACAGTAAAGATTTTGTCCGTCAGGGT	4573	ID	AX07550	standard; DNA; 5231 BP.	
DB	4233	TTTGCAAGCTGGGTATCAGTGAATTTAAACAGTAAAGATTTTGTCCGTCAGGGT	4292	XX	AX07550;		
QY	4574	GGTTTAAATTCAGAAAAAGAGAGCGAACGTCACCGTGTTCATTTGTGCAAGTGAAG	4633	AC	08-JUN-1999	(first entry)	
DB	4293	GGTTTAAATTCAGAAAAAGAGAGCGAACGTCACCGTGTTCATTTGTGCAAGTGAAG	4352	XX			
QY	4634	AGATTAAATGCTTATTTAGCGAAAAAGCGATGATTAACAGCCTTATTAGTGAAGA	4693	DT			
DB	4353	AGATTAAATGCTTATTTAGCGAAAAAGCGATGATTAACAGCCTTATTAGTGAAGA	4412	XX			
QY	4694	CAAAAAAGAGATTAGAGAGTGTAGGCATTTCTGAAACGACATTTAGATAAATTTGCTGA	4753	XX			
DB	4413	CAAAAAAGAGATTAGAGAGTGTAGGCATTTCTGAAACGACATTTAGATAAATTTGCTGA	4472	XX			
QY	4754	AGGTACTGAAGCGAATCAGGAAATTTCTTTAGATTAAACCCAGGAAGAAATCGTGCA	4813	XX			
DB	4473	AGGTACTGAAGCGAATCAGGAAATTTCTTTAGATTAAACCCAGGAAGAAATCGTGCA	4532	XX			
QY	4814	TTCAACTTGTAGTGTAAATCAATTTGTTGCTATCGATCAATTAAGTAAAGAAAGAA	4873	XX			
DB	4533	TTCAACTTGTAGTGTAAATCAATTTGTTGCTATCGATCAATTAAGTAAAGAAAGAA	4592	XX			

pTREX1A mutant beta-toxin expression cassette.  
Beta-toxin; cpb; vaccine; infection; protection; retained immunogenicity;  
haemorrhagic enteritis; necrotic enteritis; enterotoxemia;  
lamb dysentery; ds.

Synthetic.

EP892054-A1.

20-JAN-1999.

17-JUN-1998; 98BP-00202032.

20-JUN-1997; 97EP-00201888.

(ALKU ) AKZO NOBEL NV.

Sergers RPAM, Waterfield NR, Frandsen PL, Wells JM;



DR WPI; 1999-083571/08.  
XX New detoxified derivative of Clostridium perfringens beta-toxin - with  
PT retained immunogenicity, useful as a vaccine to protect against  
PT Clostridium perfringens infection.  
XX  
PS Example; Fig 1b; 30pp; English.  
XX  
CC The sequence is that of the pTREX1A expression cassette which was used in  
CC the construction of a detoxified immunogenic derivative of Clostridium  
CC perfringens beta-toxin. Such a derivative is useful as a vaccine to  
CC induce an immune response and protect against Clostridium perfringens  
CC infection in man, pig, lamb, sheep, goat, calf and bird, which causes  
CC haemorrhagic enteritis, necrotic enteritis, enterotoxemia and lamb  
CC dysentery. The new beta-toxin has been detoxified without impairing the  
CC immunogenicity, unlike prior art vaccines which used chemical and  
CC formalin detoxifying methods. Production of the beta-toxin in gram  
CC positive bacteria prevents prior art difficulties of isolation from  
CC dangerous Clostridium perfringens, and purification from the Clostridium  
CC perfringens spores  
XX  
SQ Sequence 5231 BP; 1930 A; 793 C; 985 G; 1523 T; 0 U; 0 Other;  
  
Query Match 90.0%; Score 4944.6; DB 2; Length 5231;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 4975; Conservative 0; Mismatches 4; Indels 18; Gaps 1;  
  
501 TAGTAGATCCGGTCTTAACAAAGCCGAAAGAGCTGAGTTGGCTGCGCCACCGCTG 560  
DB TCGCGGATCCGGTCTTAACAAAGCCGAAAGAGCTGAGTTGGCTGCGCCACCGCTG 312  
  
561 AGCAATAACTAGCATACCCCTTGGGCGCTTAAACGGGCTCTTAAAGGGGTTTTTCTGCA 620  
DB AGCAATAACTAGCATACCCCTTGGGCGCTTAAACGGGCTCTTAAAGGGGTTTTTCTGCA 372  
  
621 AAGGAGAACTATATCCGATGACCTGCAGGAGAGCTCTAGATCATAGCATTTTGAAG 680  
DB AAGGAGAACTATATCCG-----ACTAGATCATAGCATTTTGAAG 414  
  
681 TGGCAACAGATAAAAAAGCAGTTTAAATTTCTGCTGAATCTTTTAAACAGCAATA 740  
DB TGGCAACAGATAAAAAAGCAGTTTAAATTTCTGCTGAATCTTTTAAACAGCAATA 474  
  
741 CAATCATTTGCGAACAAGATAGCGACAGAGAGCGGAAACATTTGCTGCTGATCATTC 800  
DB CAATCATTTGCGAACAAGATAGCGACAGAGAGCGGAAACATTTGCTGCTGATCATTC 534  
  
801 ATAAAGCAATGCTTTCTTAAAGATAAAGATATTAAGAGCTATGATCAATAGTTAG 860  
DB ATAAAGCAATGCTTTCTTAAAGATAAAGATATTAAGAGCTATGATCAATAGTTAG 594  
  
861 AAAAAAGATGATCCGTAGCGGTTTTCAAAATTTGCAACAGGAATGAATTTACTATCCCT 920  
DB AAAAAAGATGATCCGTAGCGGTTTTCAAAATTTGCAACAGGAATGAATTTACTATCCCT 654  
  
921 TTTATCAAGAGCGCAAAAGAAAGAAATGATACCAATCAGTGCAGAAAGATAT 980  
DB TTTATCAAGAGCGCAAAAGAAAGAAATGATACCAATCAGTGCAGAAAGATAT 714  
  
981 AATGGGAGATAAGACGGTTCGTGCTGCTGACCTTGCACCATATCATATAAATCGAAT 1040  
DB AATGGGAGATAAGACGGTTCGTGCTGCTGACCTTGCACCATATCATATAAATCGAAT 774  
  
1041 AGCAAGAAATGGCGGAAAGCTAAAGAGTTATGGAATAAGACTTAGAAGCAACTTAA 1100  
DB AGCAAGAAATGGCGGAAAGCTAAAGAGTTATGGAATAAGACTTAGAAGCAACTTAA 834  
  
1101 GAGTGTGTTGATAGTCAGTATCTTAAATTTTGTATATAGGAATGAATTTAAATTAG 1160  
DB GAGTGTGTTGATAGTCAGTATCTTAAATTTTGTATATAGGAATGAATTTAAATTAG 894  
  
1161 ATGCTAAAAATTTGTAATTAAGAGAGTGAATTAATGAACAAAAATATAAATATCTTC 1220  
DB ATGCTAAAAATTTGTAATTAAGAGAGTGAATTAATGAACAAAAATATAAATATCTTC

895 ATGCTAAAAATTTGTAATTAAGAGAGTGAATTAATGAACAAAAATATAAATATCTTC 954  
1221 AAAAATCTTTTAAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAG 1280  
DB AAAAATCTTTTAAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAG 1014  
1281 AAAACGATACCGTTTACGAAATTTGAAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGT 1340  
DB AAAACGATACCGTTTACGAAATTTGAAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGT 1074  
1341 AAAAAGTAAACAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGT 1400  
DB AAAAAGTAAACAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGT 1134  
1401 AAAAATTAACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1460  
DB AAAAATTAACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1194  
1461 TCCCTTAAACAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGT 1520  
DB TCCCTTAAACAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGT 1254  
1521 TTAATTAACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1580  
DB TTAATTAACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1314  
1581 GATTTCTAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1640  
DB GATTTCTAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1374  
1641 TCTCGATTCAGCAATTTGCTTAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1700  
DB TCTCGATTCAGCAATTTGCTTAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1434  
1701 ACAGTGTCTTAATAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1760  
DB ACAGTGTCTTAATAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1494  
1761 TATATACGCTTCTTTTCAAAATTTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1820  
DB TATATACGCTTCTTTTCAAAATTTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1554  
1821 ATCAATTTTCAATCAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1880  
DB ATCAATTTTCAATCAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1614  
1881 AGCAAGTATTCCTTATTTTAAATTTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1940  
DB AGCAAGTATTCCTTATTTTAAATTTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1674  
1941 AGTGTCTTTTGTAAATTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2000  
DB AGTGTCTTTTGTAAATTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1734  
2001 GGTATATCTGACAGCTTCCAGAGGTTAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2060  
DB GGTATATCTGACAGCTTCCAGAGGTTAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1794  
2061 AGCTCGGATCATATGCAAGCAAAATTAACCTCGCAACAGCACTTTGGAGAAATGGGACGA 2120  
DB AGCTCGGATCATATGCAAGCAAAATTAACCTCGCAACAGCACTTTGGAGAAATGGGACGA 1854  
2121 TCGAGAAAACCTCTTTTACGCTGAGTGAATATATTAATAAGCCGTGAAGAGAGCGGTTTC 2180  
DB TCGAGAAAACCTCTTTTACGCTGAGTGAATATATTAATAAGCCGTGAAGAGAGCGGTTTC 1914  
2181 AAAAAGGTTTTTAAATAAGGAGAGCAATCAATGATTTAGTGAAGTGAAGTGAAGTGAAG 2240  
DB AAAAAGGTTTTTAAATAAGGAGAGCAATCAATGATTTAGTGAAGTGAAGTGAAGTGAAG 1974  
2241 AACGTGAGAGATTTAGAGAGAGTGTCTCTCCAAGACCAAGTGAAGAGTGAAGTGAAGTGA 2300  
DB AACGTGAGAGATTTAGAGAGAGTGTCTCTCCAAGACCAAGTGAAGAGTGAAGTGAAGTGA 2034



```
Db 4195 AAGTGCCTATTCAGAAACTATCAAGGGCTAATAGGGAATACATTAACCATTCCTTGCAA 4254
Qy 4521 AGCTTGGGTATCAAGTGAATTAACCACTAAAGATTTATTTGTCGGTCAAGGGTGGTTAA 4580
Db 4255 AGCTTGGGTATCAAGTGAATTAACCACTAAAGATTTATTTGTCGGTCAAGGGTGGTTAA 4314
Qy 4581 ATTCAAGAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4640
Db 4315 ATTCAAGAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4374
Qy 4641 AATGGCTTATATTAGCGAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4700
Db 4375 AATGGCTTATATTAGCGAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4434
Qy 4701 AGAGATTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4760
Db 4435 AGAGATTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4494
Qy 4761 GAAGCGCAATCAGGAAATTTCTTTTAAAGATTAACCAAGAGAGAGAGAGAGAGAGAGAGAG 4820
Db 4495 GAAGCGCAATCAGGAAATTTCTTTTAAAGATTAACCAAGAGAGAGAGAGAGAGAGAGAGAG 4554
Qy 4821 TGCTAGTGTAAATCATCTGTTGCTATCGATCATTAAGATTAAGATTAAGATTAAGATTAAG 4880
Db 4555 TGCTAGTGTAAATCATCTGTTGCTATCGATCATTAAGATTAAGATTAAGATTAAGATTAAG 4614
Qy 4881 CTATATAAGCGCTGCAAAATTTCTTTTAAAGATTAACCAAGAGAGAGAGAGAGAGAGAGAGAG 4940
Db 4615 CTATATAAGCGCTGCAAAATTTCTTTTAAAGATTAACCAAGAGAGAGAGAGAGAGAGAGAGAG 4674
Qy 4941 AACAAGCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5000
Db 4675 AACAAGCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4734
Qy 5001 AGGCTGAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5060
Db 4735 AGGCTGAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4794
Qy 5061 TTTCTTTGCTGTTAGCGAATGATTAGCGAATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5120
Db 4795 TTTCTTTGCTGTTAGCGAATGATTAGCGAATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4854
Qy 5121 TTAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5180
Db 4855 TTAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4914
Qy 5181 AGGGAAGGCTACTAAACCTGAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5240
Db 4915 AGGGAAGGCTACTAAACCTGAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4974
Qy 5241 ATTTCTATCTTTTATAGTCAATAGAGATATCTTATTTGCTATATAAATCTATTAGCAG 5300
Db 4975 ATTTCTATCTTTTATAGTCAATAGAGATATCTTATTTGCTATATAAATCTATTAGCAG 5034
Qy 5301 CATAATAGATTTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5360
Db 5035 CATAATAGATTTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5094
Qy 5361 AGAAATATTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5420
Db 5095 AGAAATATTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5154
Qy 5421 TAACTAAAGTAGTGAATTTTGAATTTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5480
Db 5155 TAACTAAAGTAGTGAATTTTGAATTTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5214
Qy 5481 GTCAAGTAGTGAATTAATA 5497
Db 5215 GTCAAGTAGTGAATTAATA 5231
```

RESULT 9

```
AAI3224/c
ID AAI3224 standard; DNA; 3765 BP.
XX
XX AAI3224;
XX AC
XX DT 19-MAR-1999 (first entry)
XX
DE Enterococcus faecalis genome contig SEQ ID NO:287.
XX
XX Enterococcus faecalis; contig; detection; Enterococcal infection;
KW vaccine; attenuation; computer readable medium; ds.
XX
XX Enterococcus faecalis.
OS
XX WO9850555-A2.
XX
XX 12-NOV-1998.
XX
XX 04-MAY-1998; 98WO-US008985.
XX
XX 06-MAY-1997; 97US-0044031P.
XX 16-MAY-1997; 97US-0046655P.
XX 14-NOV-1997; 97US-0066009P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Kunsch CA, Dillon PJ, Barash SC;
XX WPI; 1999-045171/04.
XX
XX New isolated Enterococcus faecalis polynucleotides and polypeptides -
used to develop products for the detection of Enterococcus and for use in
vaccines for prevention or attenuation of Enterococcus infection.
XX
XX Claim 1; Page 1316-1317; 2084pp; English.
XX
XX A computer readable medium has been developed which has recorded on it
982 nucleotide sequences isolated from the Enterococcus faecalis genome.
XX AAI3224 to AAI3229 represent these nucleotide sequences which are
XX primary nucleotide sequences, also known as contigs. The computer-based
XX system can identify fragments of the Enterococcus faecalis genome with
XX commercial importance. The products can be used to detect the presence of
XX Enterococcus faecalis in samples. They can also be used for diagnosing
XX Enterococcal infection in an animal and monitoring progression of
XX disease, and for identifying agents which can be used to modulate the
XX growth or pathogenicity of Enterococcus faecalis, or another related
XX organism, in vivo or in vitro. In particular the polypeptides encoded by
XX the Enterococcus faecalis nucleotide sequences can be used in vaccines to
XX prevent or attenuate an Enterococcal infection
XX
XX Sequence 3765 BP; 1172 A; 681 C; 638 G; 1263 T; 0 U; 11 Other;
Query Match 26.1%; Score 1434.8; DB 2; Length 3765;
Best Local Similarity 96.6%; Pred. No. 3.6e-277;
Matches 1485; Conservative 2; Mismatches 49; Indels 2; Gaps 2;
Qy 3488 TAGGATTTATAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3547
Db 2075 TAAATTTTGGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2016
Qy 3548 TTCATGCGGCTTGTAAAGTACAAATTTAAATAGTAAATTCGCTCAATCTACCA 3607
Db 2015 TTCATGCGGCTTGTAAAGTACAAATTTAAATAGTAAATTCGCTCAATCTACCA 1956
Qy 3608 AGCAGGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3667
Db 1955 AGCAGGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1896
Qy 3668 GTCGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3727
Db 1895 GACGTGGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1836
Qy 3728 ATTGGACACCAACGTTTATCGTTATGGAACGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 3787
```

1835 ATTGGACACCAACGTTTATCGTTATCGTATCGATGACGAGCAAAACCGTTTACACGA 1776  
 3788 AAGGACATCTGAAAACAAATTAAGACAAATCAATCTCTTTATTGATTTTGTATTC 3847  
 1775 AAGGACATCTGAAAACAAATTAAGACAAATCAATCTCTTTATTGATTTTGTATTC 1716  
 3848 ACACGCAAAAGAAACTATTTTCAGCAAGCGATATTTTAAACACGCTTATGATTTAGGT 3907  
 1715 ACACGCAAAAGAAACTATTTTCAGCAAGCGATATTTTAAACACGCTTATGATTTAGGT 1656  
 3908 TTATGCTCTATGATTTATCAATCTGATTAAGGTTATCAAGCATATTTTGTATTTAGAA 3967  
 1655 TTATGCTCTATGATTTATCAATCTGATTAAGGTTATCAAGCATATTTTGTATTTAGAA 1596  
 3968 CGCCAGTCTATGCTATTCAGCAATCAAAATCAGAAATTTAAATCTGTCAAAGCAGCCAAATATTT 4027  
 1595 CGCCAGT-TATGCTATTCAGCAATCAAAATCAGAAATTTAAATCTGTCAAAGCAGCC-AAATATCT 1538  
 4028 CGCAAAATATCCGAGATATTTTGGAAAGCTTTTCCAGTGTATCTAACGTTATCATTT 4087  
 1537 CGCAAAATATCCGAGATATTTTGGAAAGCTTTTCCAGTGTATCTAACGTTATCATTT 1478  
 4088 TTGATTTGCTCGCATACCAAGACGACATGAGAAATTTTGTATCTTAATACCGTT 4147  
 1477 TTGGATGCTCGTATACCAAGACGACATGAGAAATTTTGTATCTTAATACCGTT 1418  
 4148 ATTCTTTCAAAGATGCGAAGTTGCTTTTCAAACAAACAGATAATTAAGGCTTTACTC 4207  
 1417 ATTCTTTCAAAGATGCGAAGTTGCTTTTCAAACAAACAGATAATTAAGGCTTTACTC 1358  
 4208 GTTCAAGTCTAACGTTTAAAGCGGTACAGAGCGCAAAACAGATGATGAACTCTGGT 4267  
 1357 GTTCAAGTCTAACGTTTAAAGCGGTACAGAGCGCAAAACAGATGATGAACTCTGGT 1298  
 4268 TTAATCTCTTATGTCAGAAACGAAATTTTTCAGGAGAAAGGTTTAAATAGGCGTAA 4327  
 1297 TTAATCTCTTATGTCAGAAACGAAATTTTTCAGGAGAAAGGTTTAAATAGGCGTAA 1238  
 4328 AGTCATGTTTACCTCTCTTTAGCTTATCTTATGCTAGGCTATCTCAATCGAACGTCG 4387  
 1237 AGTCATGTTTACCTCTCTTTAGCTTATCTTATGCTAGGCTATCTCAATCGAACGTCG 1178  
 4388 AATATAATATGTTGAGTTTAAATATCGATTAATCAACCTTAAAGAGAAAGAAAGTAA 4447  
 1177 AATATAATATGTTGAGTTTAAATATCGATTAATCAACCTTAAAGAGAAAGAAAGTAA 1118  
 4448 TCAAAATTTAGAGTGCCTTATCAGAAATCTATCAAGGGCTTAAAGGATACATTA 4507  
 1117 TTAACCTTTGTAAGAGTGCCTTATCAGAAATCTATCAAGGGCTTAAAGGATACATTA 1058  
 4508 CCATCTTTTGAAGCTTTGGGTATCAAGTATTAACCAAGTAAAGATTTATTTGTCGCTC 4567  
 1057 CCATCTTTTGAAGCTTTGGGTATCAAGTATTAACCAAGTAAAGATTTATTTGTCGCTC 998  
 4568 AAGGCTGTTTAAATTCAGAAACAAAGAGCAAGCGTCAACGTTGATTTGTCAGAT 4627  
 997 AAGGCTGTTTAAATTCAGAAACAAAGAGCAAGCGTCAACGTTGATTTGTCAGAT 938  
 4628 GGAAGAGATTTATGCTTATATACGAAAGAGGATGATGATGATGATGATGATGATGATGAT 4687  
 937 GGAAGAGATTTATGCTTATATACGAAAGAGGATGATGATGATGATGATGATGATGATGAT 878  
 4688 TGACGACCAAAAGAGATTTAGAGAGTGTAGGCAATTCCTGAACGCAATTAATTAAT 4747  
 877 TGACGACCAAAAGAGATTTAGAGAGTGTAGGCAATTCCTGAACGCAATTAATTAAT 818  
 4748 TGCTGAAGTACTGAAGGCAATCAGGAATTTCTTAAAGTAAACCAAGCAAGAAATG 4807  
 817 TGCTGAAGTACTGAAGGCAATCAGGAATTTCTTAAAGTAAACCAAGCAAGAAATG 758  
 4808 GTGGCAATCAATCTGCTAGTGTAAATCATTTGCTATCGATCAATTAAGTAAAGAAAG 4867

Db 757 GTGGCAATCAATCTGCTAGTGTAAATCATTTGTTGCTATCGATCAATTAAGTAAAGAAAG 698  
 QY 4868 AAGAAAAAGAAAGTATATAAAGCGCTGACAAATTTCTTTGACTTAGAGCATACATCA 4927  
 Db 697 AAGAAAAAGAAAGTATATAAAGCGCTGTCAGAGTTTTTTGACTTAGAACATACATCA 638  
 QY 4928 TTCAGAGACTTTTAAACAGCTAGCAGAACGCCCTTAAACGAGACACAACTCGATTTGT 4987  
 Db 637 TTCAGAGACTTTTAAACAGCTAGCAGAACGCCCTTAAACGAGACACAACTCGATTTGT 578  
 QY 4988 TTAGCTATGATACAGGCTGAAATAAACCCTGACTAT 5025  
 Db 577 TTAGCTATGATACAGGCTGAAATAAAGTAAAGTCGGTAT 540  
 RESULT 10  
 ABS99019/c  
 ID ABS99019 standard; DNA; 3765 BP.  
 XX  
 AC ABS99019;  
 XX  
 DT 18-DEC-2002 (first entry)  
 XX  
 DE Enterococcus faecalis contig sequence #287.  
 XX  
 KW Computer readable medium; Enterococcus faecalis; microbe; growth;  
 KW pathogenicity; vaccine; resistance; Enterococcal infection; commercial;  
 KW therapeutic; industrial; fermenting; sugar source; metabolite; vaccine;  
 KW biotech technology; antibacterial; modulator of nucleic acid expression;  
 KW contig; ds.  
 XX  
 OS Enterococcus faecalis.  
 XX  
 PN US2002120116-A1.  
 XX  
 PD 29-AUG-2002.  
 XX  
 PF 04-MAY-1998; 98US-00070927.  
 XX  
 PR 04-MAY-1998; 98US-00070927.  
 XX  
 PA (KUNS/) KUNSCH C A.  
 PA (DILL/) DILLON P J.  
 PA (BARA/) BARASH S.  
 XX  
 PI Kunsch CA, Dillon PJ, Barash S;  
 XX  
 PS WPI; 2002-750065/81.  
 CC  
 CC Computer readable medium having recorded on it a Enterococcus faecalis  
 CC nucleotide sequence useful for detecting diseases related to Enterococcus  
 CC infections in animals.  
 CC  
 CC Claim 1; Page; 119pp; English.  
 CC  
 CC The present invention relates to a new computer readable medium with an  
 CC Enterococcus faecalis nucleotide sequence. The invention is useful to  
 CC diagnose the presence of E.faecalis in a sample or determining the  
 CC presence of a specific microbe in a sample. The invention is also useful  
 CC for modulating the growth or pathogenicity of E.faecalis, in a vaccine to  
 CC confer resistance to Enterococcal infection, for commercial, therapeutic  
 CC and industrial purposes, and for fermenting a particular sugar source or  
 CC to produce a particular metabolite. The invention is useful for detecting  
 CC diseases related to Enterococcus infections in animals, and for detecting  
 CC E.faecalis using biotech technology. The present nucleic acid sequence  
 CC represents an Enterococcus faecalis contig DNA sequence of the invention.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification but was obtained in electronic format directly from USPTO  
 CC at http.sequdata.uspto.gov  
 XX  
 SQ Sequence 3765 BP; 1172 A; 681 C; 638 G; 1263 T; 0 U; 11 Other;  
 Query Match 26.1%; Score 1434.8; DB 6; Length 3765;

1057	DB	CGATTCCTTTGCAAGAGCTTCGGGTATCAAGTGATGATTTAAACCAAGTAAAGATTTATTTTGTCCGTC	998
4568	QY	AAGGGTGGTTTTAAATTCAGAAAAAAGAGCGAACGCTCAACCGTGTTCATTTTGTCAGAAT	4627
997	DB	AAGGGTGGTTTTAAATTCAGAAAAAAGAGTGAACGCTCAACCGTGTTCATTTTGTCAGAAT	938
4628	QY	CGNAGCAAGATTTAATGGCTTATATTAGCGAAAAAAGCGATGTATACAAGCCTTATTATTAG	4687
937	DB	GGAAAGAAAGATTTAATGGCTTATATTAGCGAAAAAAGTGATGTATACAAGCCTTATTATTGG	878
4688	QY	TGACGACCAAAAAAGAGATTAGAGAAGTGCTAGGCATTCCTGAAACGGACATTAGATAAAT	4747
877	DB	TGACGACTAAAAAGAGATTAGAGAAGCGCTAGGCATTCCTGAGCGGACATTAGATAAAT	818
4748	QY	TGCTGAAGTACTGAAGGCGAATCAGGAATTTTCTTTAAGATTAAACCGAGGAAGAAATG	4807
817	DB	TGCTGAAGTACTGAAGGCGAATCAGGAATTTTCTTTAAGATTAAACCGAGGAAGAAATG	758
4808	QY	GTGGCATTCCAACTTGCTAGTGTAAATCATTTGCTATCGATCATTTAAAGTAAAAAAG	4867
757	DB	GTGGCATTCCAACTTGCTAGTGTAAATCATTTGCTATCGATCATTTAAAGTAAAAAAG	698
4868	QY	AGAAAAAGAAAGCTATATAAAGGCGCTGACAAATTCCTTTGACTTGAAGCATACATTC	4927
697	DB	AGAAAAAGAAAGCTATATAAAGGCGCTGACAGATTTTTCGACTTGAAGCATACATTC	638
4928	QY	TTCAGAGACTTTAAACAAGCTAGCAGACGCGCTTAAACGGACACACAACTCGATTGT	4987
637	DB	TTCAGAGACTTTAAACAAGCTAGCAGACGCGCTTAAACGGACACACAACTCGATTGT	578
4988	QY	TTAGCTATGATACAGCTGAAAAATTAACCCGCACTAT	5025
577	DB	TTAGCTATGATACAGCTGAAATAGATAAAGTCCGSTAT	540
RESULT 11			
AAAX90954			
ID	AAAX90954 standard; DNA; 10323 BP.		
AC	AAAX90954;		
XX	17-JAN-2000 (first entry)		
DT	Plant transformation binary vector, plasmid pDAB1542.		
DE	Palmitate-CoA delta-9 desaturase; PCA; palmitic acid; desaturate; palm;		
KW	modification; saturate oil; fatty acid; maize; soybean; construct;		
KW	promoter regulatory element; ubiquitin promoter; transfection; coconut;		
KW	maize globulin promoter; transcriptional terminator sequence; sunflower;		
KW	seed embryo cell; host plant; safflower; transgenic plant; binary vector;		
KW	sense orientation; plasmid pDAB1542; high level expression; ds.		
XX	Agrobacterium tumefaciens.		
OS	Escherichia coli.		
OS	Cauliflower mosaic virus.		
XX	Synthetic.		
Key	Location/Qualifiers		
misc_feature	1..579		
FT	/*tag= a		
FT	/note= "Correspond to bases 602-1184 of plasmid pTI-15955		
FT	derived from A.tumefaciens"		
FT	304..327		
FT	/*tag= b		
FT	/label= T DNA Border A sequence		
FT	/note= "Derived from Agrobacterium tumefaciens"		
FT	580..597		
FT	/*tag= c		
FT	/note= "Remnants of cloning manipulations"		
transposon	598..2027		
FT	/*tag= d		
FT	/label= Transposon Tn903		
FT	/note= "Correspond to bases 835-2264 of Tn903 derived		

```
FT FT from E.coli"
FT FT 925..1740
FT FT /tag= e
FT FT /product= "Neomycin phosphotransferase I protein"
FT FT /note= "Derived from E.coli transposon Tn903"
FT FT replace(1230, C)
FT FT /tag= f
FT FT Tn903 and destroys SmaI recognition site"
FT FT /note= "Corresponds to base 1467 of E.coli transposon
FT FT Tn903 and destroys SmaI recognition site"
FT FT replace(1477, C)
FT FT /tag= g
FT FT /note= "Corresponds to base 1714 of E.coli transposon
FT FT Tn903 and destroys HindIII recognition site"
FT FT 2028..2062
FT FT /tag= h
FT FT /note= "Remnants of cloning manipulations"
FT FT complement(2063..2080)
FT FT /tag= i
FT FT /label= Transposon Tn5
FT FT /note= "Correspond to bases 2519-2536 of Tn5 derived from
FT FT E.coli"
FT FT 2081..2793
FT FT /tag= j
FT FT /note= "Correspond to bases 21728-22440 of plasmid pri-
FT FT 15955 derived from A.tumefaciens"
FT FT complement(2794..3772)
FT FT /tag= k
FT FT /note= "Correspond to bases 1540-2518 of Tn5 transposon
FT FT derived from E.coli"
FT FT complement(2967..3761)
FT FT /tag= l
FT FT /product= "Neomycin phosphotransferase II protein"
FT FT /note= "Derived from E.coli transposon Tn5"
FT FT replace(3764, G)
FT FT /tag= m
FT FT /note= "Corresponds to base 1532 of transposon Tn5
FT FT derived from E.coli"
FT FT replace(3768, C)
FT FT /tag= n
FT FT /note= "Corresponds to base 1536 of transposon Tn5
FT FT derived from E.coli to create a Bam HI site"
FT FT 3773..3784
FT FT /tag= o
FT FT /note= "Remnants of cloning manipulations"
FT FT 3785..4174
FT FT /tag= p
FT FT /label= 19S promoter sequence
FT FT /note= "Correspond to bases 5376-5765 of Cabbs strain of
FT FT Cauliflower Mosaic virus"
FT FT 4175..4272
FT FT /tag= q
FT FT /label= Multiple cloning site
FT FT /note= "Comprises restriction sites for Bgl I, Asc I, Sma
FT FT I, Srf I, Pme I, Not I and Pac I"
FT FT 4273..4624
FT FT /tag= r
FT FT /note= "Correspond to bases 13926-14277 of plasmid pri-
FT FT 15955 derived from A.tumefaciens"
FT FT 4407..4432
FT FT /tag= s
FT FT /label= T-DNA Border B sequence
FT FT /note= "Derived from Agrobacterium tumefaciens"
FT FT 4445..4468
FT FT /tag= t
FT FT /label= Overdrive sequence
FT FT /note= "Derived from plasmid pri-15955 of A.tumefaciens"
FT FT 4631..5433
FT FT /tag= u
FT FT /label= Erythromycin resistance determinant
FT FT /note= "Derived from plasmid pR29"
FT FT 5434..5828
FT FT /tag= v
FT FT /label= STREAMW1_sequence
FT FT
FT CDS
FT FT 925..1740
FT FT /tag= e
FT FT /product= "Neomycin phosphotransferase I protein"
FT FT /note= "Derived from E.coli transposon Tn903"
FT FT replace(1230, C)
FT FT /tag= f
FT FT Tn903 and destroys SmaI recognition site"
FT FT /note= "Corresponds to base 1467 of E.coli transposon
FT FT Tn903 and destroys SmaI recognition site"
FT FT replace(1477, C)
FT FT /tag= g
FT FT /note= "Corresponds to base 1714 of E.coli transposon
FT FT Tn903 and destroys HindIII recognition site"
FT FT 2028..2062
FT FT /tag= h
FT FT /note= "Remnants of cloning manipulations"
FT FT complement(2063..2080)
FT FT /tag= i
FT FT /label= Transposon Tn5
FT FT /note= "Correspond to bases 2519-2536 of Tn5 derived from
FT FT E.coli"
FT FT 2081..2793
FT FT /tag= j
FT FT /note= "Correspond to bases 21728-22440 of plasmid pri-
FT FT 15955 derived from A.tumefaciens"
FT FT complement(2794..3772)
FT FT /tag= k
FT FT /note= "Correspond to bases 1540-2518 of Tn5 transposon
FT FT derived from E.coli"
FT FT complement(2967..3761)
FT FT /tag= l
FT FT /product= "Neomycin phosphotransferase II protein"
FT FT /note= "Derived from E.coli transposon Tn5"
FT FT replace(3764, G)
FT FT /tag= m
FT FT /note= "Corresponds to base 1532 of transposon Tn5
FT FT derived from E.coli"
FT FT replace(3768, C)
FT FT /tag= n
FT FT /note= "Corresponds to base 1536 of transposon Tn5
FT FT derived from E.coli to create a Bam HI site"
FT FT 3773..3784
FT FT /tag= o
FT FT /note= "Remnants of cloning manipulations"
FT FT 3785..4174
FT FT /tag= p
FT FT /label= 19S promoter sequence
FT FT /note= "Correspond to bases 5376-5765 of Cabbs strain of
FT FT Cauliflower Mosaic virus"
FT FT 4175..4272
FT FT /tag= q
FT FT /label= Multiple cloning site
FT FT /note= "Comprises restriction sites for Bgl I, Asc I, Sma
FT FT I, Srf I, Pme I, Not I and Pac I"
FT FT 4273..4624
FT FT /tag= r
FT FT /note= "Correspond to bases 13926-14277 of plasmid pri-
FT FT 15955 derived from A.tumefaciens"
FT FT 4407..4432
FT FT /tag= s
FT FT /label= T-DNA Border B sequence
FT FT /note= "Derived from Agrobacterium tumefaciens"
FT FT 4445..4468
FT FT /tag= t
FT FT /label= Overdrive sequence
FT FT /note= "Derived from plasmid pri-15955 of A.tumefaciens"
FT FT 4631..5433
FT FT /tag= u
FT FT /label= Erythromycin resistance determinant
FT FT /note= "Derived from plasmid pR29"
FT FT 5434..5828
FT FT /tag= v
FT FT /label= STREAMW1_sequence
FT FT
FT misc_feature
FT FT 925..1740
FT FT /tag= w
FT FT /label= EHERMAM_sequence
FT FT 5546..6283
FT FT /tag= x
FT FT /product= "Adenine methylase protein"
FT FT /note= "Putative"
FT FT replace(5586, A)
FT FT /tag= y
FT FT /note= "Derived from EHERMAM sequence"
FT FT replace(5927, A)
FT FT /tag= z
FT FT /note= "Derived from EHERMAM sequence"
FT FT replace(5930, A)
FT FT /tag= aa
FT FT /note= "Derived from EHERMAM sequence"
FT FT replace(5931, A)
FT FT /tag= ab
FT FT /note= "Derived from EHERMAM sequence"
FT FT replace(5933, A)
FT FT /tag= ac
FT FT /note= "Derived from EHERMAM sequence"
FT FT replace(5934, A)
FT FT /tag= ad
FT FT /note= "Derived from EHERMAM sequence"
FT FT replace(5935, A)
FT FT /tag= ae
FT FT /note= "Derived from EHERMAM sequence"
FT FT replace(5936, A)
FT FT /tag= af
FT FT /note= "Derived from EHERMAM sequence"
FT FT replace(5938, A)
FT FT /tag= ag
FT FT /note= "Derived from EHERMAM sequence"
FT FT 5943..6448
FT FT /tag= ah
FT FT /label= STREAMM2 sequence
FT FT /note= "Correspond to bases 1-506"
FT FT 6448..8866
FT FT /tag= ai
FT FT /label= Plasmid RK2
FT FT /note= "Correspond to bases 15435-17853"
FT FT 6574
FT FT /tag= aj
FT FT /note= "Additional 'T' nucleotide"
FT FT replace(6651, C)
FT FT /tag= al
FT FT /note= "Derived from plasmid RK2 sequence"
FT FT 6905
FT FT /tag= ak
FT FT /note= "Additional 'C' nucleotide"
FT FT replace(7446, C)
FT FT /tag= am
FT FT /note= "Derived from plasmid RK2 sequence"
FT FT replace(7461, C)
FT FT /tag= an
FT FT /note= "Derived from plasmid RK2 sequence"
FT FT replace(7479, G)
FT FT /tag= ao
FT FT /note= "Derived from plasmid RK2 sequence"
FT FT replace(7494, C)
FT FT /tag= ap
FT FT /note= "Derived from plasmid RK2 sequence"
FT FT 8861..9602
FT FT /tag= aq
FT FT /label= Plasmid RK2
FT FT /note= "Correspond to bases 50632-51373"
FT FT 9603..9613
FT FT /tag= ar
FT FT /note= "Remnants of cloning manipulations"
FT FT complement(9614..10322)
FT FT /tag= as
```



	/label= Plasmid_RK2	
FT	/note= "Correspond to bases 12109-12817"	
FT	replace(9742, A)	
FT	/*tag= at	
FT	/note= "Derived from plasmid RK2 sequence"	
FT	replace(10024, A)	
FT	/*tag= au	
FT	/note= "Derived from plasmid RK2 sequence"	
FT	10191	
FT	/*tag= av	
FT	/note= "Additional 'T' nucleotide"	
FT	10323	
FT	/*tag= aw	
FT	/note= "Remnants of cloning manipulations"	
XX		
PN	W09950430-A2.	
XX		
PD	07-OCT-1999.	
	Query Match	20.2%; Score 1111; DB 2; Length 10323;
	Best Local Similarity	100.0%; Pred. No. 2.8e-212; Gaps 0;
	Matches 1111; Conservative 0; Mismatches 0; Indels 0;	
Qy	936 AAAGAAGAAAAACGAAATGATACACCAATCAGTCGCAAAAAGAGATAAATGGAGATAAGAC	995
Dd	5286 AAAAGAAGAAAAACGAAATGATACACCAATCAGTCGCAAAAAGAGATAAATGGAGATAAGAC	5345
Qy	996 GGTTCGGTGTCGTGCTGACTTGACCCATATCATAAAAATCGAAACACGAAGAATGCCG	1055
Dd	5346 GGTTTCGTGTTGTCGTGCTGACTTGACCCATATCATAAAAATCGAAACACGAAGAATGCCG	5405
Qy	1056 AAACGCTAAAAAGAACGTTATGGAAATTAAGCATTAAGAAGCAAACCTTAAGAGTGTGTTGATGT	1115
Dd	5406 AAACGCTAAAAAGAACGTTATGGAAATTAAGCATTAAGAAGCAAACCTTAAGAGTGTGTTGATGT	5465
Qy	1116 GCAGTAGTCTTAAAAATTTTGTATTAATTCGAANTTGAAGTTAAATTAGATGCTTAAAAATTTGT	1175
Dd	5466 GCAGTAGTCTTAAAAATTTTGTATTAATTCGAANTTGAAGTTAAATTAGATGCTTAAAAATTTGT	5525
Qy	1176 AATTAAAGACGAGTGATTACATGAACAAAAATATAAAAAATATCTCAAAACCTTTTAAACGA	1235
Dd	5526 AATTAAAGACGAGTGATTAATGAACAAAAATATAAAAAATATCTCAAAACCTTTTAAACGA	5585
Qy	1236 GTGAAAAAGTACTCAACCAAATATAAAAACAAATTTAAAGAAACCGCATACCGTTT	1295
Dd	5586 GTGAAAAAGTACTCAACCAAATATAAAAACAAATTTAAAGAAACCGCATACCGTTT	5645
Qy	1296 ACGBAATTGGAACAGGTAAAGGCATTTACGCGBAACCTGGCTTAAATTAAGTAAACAGG	1355
Dd	5646 ACGBAATTGGAACAGGTAAAGGCATTTACGCGBAACCTGGCTTAAATTAAGTAAACAGG	5705
Qy	1356 TAACGTCATTGTAATTAGACAGTCATCTATTCAACTTATCGTCAGAAAAATTTAAACCTGA	1415
Dd	5706 TAACGTCATTGTAATTAGACAGTCATCTATTCAACTTATCGTCAGAAAAATTTAAACCTGA	5765
Qy	1416 ATACTCGTGTCACTTTAAATTCACCAAGATATCTACAGTTTCAATTCCTCTAACCAACAGA	1475
Dd	5766 ATACTCGTGTCACTTTAAATTCACCAAGATATCTACAGTTTCAATTCCTCTAACCAACAGA	5825
Qy	1476 GGTATAAAATTTGTTGGAGTATTCCCTTACCAATTTAAGCACACAATTTATTTAAAAAAGTGG	1535
Dd	5826 GGTATAAAATTTGTTGGAGTATTCCCTTACCAATTTAAGCACACAATTTATTTAAAAAAGTGG	5885
Qy	1536 TTTTTGAAAGCCATGGCTGTGACATCTATCTGATTTGTTGAAGAAAGGATTTCTACAGCGTA	1595
Dd	5886 TTTTTGAAAGCCATGGCTGTGACATCTATCTGATTTGTTGAAGAAAGGATTTCTACAGCGTA	5945
Qy	1596 CTTTGGATATTTACCAGAACACTAGGGTTGCTTCTTGCACTCAAGTCTCGATTACAGCAAT	1655
Dd	5946 CTTTGGATATTTACCAGAACACTAGGGTTGCTTCTTGCACTCAAGTCTCGATTACAGCAAT	6005
Qy	1656 TGCTTAAAGCTGCCACGCGGAATGCTTTTCATCCTAAACCAAAAGTAAACAGTGTCTTTAATAA	1715

6006	TGCTTAAGCTGCCGACGGAAATGCTTTCATCTCTAAACCAAAGTAAACAGTGTCTTAATAA	6065
1716	AACCTACCCGCCCATACACAGATGTTCCAGATAAATATTCGAAAGCTATATACGTA	1775
6066	AACCTACCCGCCCATACACAGATGTTCCAGATAAATATTCGAAAGCTATATACGTA	6125
1776	TTTCAAAATGGGTCAATCGAGAAATATCGTCAACTGTTTACTAAAAATCAGTTTCATCAAG	1835
6126	TTTCAAAATGGGTCAATCGAGAAATATCGTCAACTGTTTACTAAAAATCAGTTTCATCAAG	6185
1836	CAATGAACAACGCGCAAAGTAAACAATTTAAGTACCGTTTACTTATGACCAAGTATTTGTCTA	1895
6186	CAATGAACAACGCGCAAAGTAAACAATTTAAGTACCGTTTACTTATGACCAAGTATTTGTCTA	6245
1896	TTTTTAATAGTTATCTATTATTTAAACGGGAGGAAATTAATTCATAGTCGCTTTTGTAAA	1955
6246	TTTTTAATAGTTATCTATTATTTAAACGGGAGGAAATTAATTCATAGTCGCTTTTGTAAA	6305
1956	TTTGGAAAGTTACACGTTACTTAAAGGGAATGTAGATAAAATTTAGGTATATCTACTGACA	2015
6306	TTTGGAAAGTTACACGTTACTTAAAGGGAATGTAGATAAAATTTAGGTATATCTACTGACA	6365
2016	GCCTTCCAAGGAGCTAAAGAGGTCCTTAGCGC	2046
6366	GCCTTCCAAGGAGCTAAAGAGGTCCTTAGCGC	6396

RESULT 12

AAD01289

ID AAD01289 standard; DNA; 17111 BP.

XX AAD01289;

XX XX

DT 12-OCT-2000 (first entry)

XX XX

DE Arabidopsis transformation binary vector, pAract2Af-bin.

KW Matrix Attachment Region; MAR; scaffold attachment region; arabidopsis; transformation binary vector; pAract2Af-bin; gene expression; transgenic organism; ds.

KW KW

XX XX

OS Synthetic.

XX XX

FH Key Location/Qualifiers

FT misc\_feature 50..712

FT FT /\*tag= a

FT FT /note= "Corresponds to nucleotides 14-676 of MAR dimer-2"

FT misc\_feature 4039..4893

FT FT /\*tag= b

FT FT /note= "Corresponds to nucleotides 1-646 of MAR dimer-1"

XX XX

PN WO200032800-A1.

XX XX

PD 08-JUN-2000.

XX XX

PF 30-NOV-1999; 99WO-US028123.

XX XX

PR 01-DEC-1998; 98US-0110437P.

XX XX

PA (DOWC ) DOW AGROSCIENCES LLC.

XX XX

PI Van Der Geest AHM, Ainley WM, Cowen NM, Welter ME, Woosley AT;

XX XX

DR WPI; 2000-412345/35.

XX XX

PT An isolated DNA molecule for use as a matrix attachment region to increase expression of genes introduced in transformed plants comprises a 298 base pair sequence described in the specification.

PT PT

XX XX

PS Example 3; Page 59-64; 73pp; English.

XX XX

CC The patent discloses a DNA molecule, useful as matrix attachment region (MAR) or scaffold attachment region, to increase the expression of genes

An isolated DNA molecule for use as a matrix attachment region to increase expression of genes introduced in transformed plants comprises a 298 base pair sequence described in the specification.

Example 3; Page 59-64; 73pp; English.

The patent discloses a DNA molecule, useful as matrix attachment region (MAR) or scaffold attachment region, to increase the expression of gene

CC	introduced in transformed plants. MARs are located in non-transcribed
CC	regions of genes and form the physical boundaries of individual DNA
CC	loops. They are rich in adenosine and thymine bases and contain certain
CC	conserved sequence elements and structural features. They are about 300-
CC	2000 bp in length. Increased levels of expression of DNA introduced into
CC	plants can be achieved by use of MAR. They can also reduce the position
CC	effect in transgenic organisms. The present DNA sequence is the
CC	Arabidopsis transformation binary vector pARCTAfl-bin, identical to the
CC	vector pACT2-bin, except that it contains a MAR dimer-2 positioned 5' to
CC	the Act2 transcription initiation region and the MAR dimer-1 positioned
CC	3' to the nos 3' UTR (untranslated region). This vector is used to test
CC	the two orientations of the artificial MAR dimer in Arabidopsis.
XX	
SQ	Sequence 17111 BP; 4720 A; 3952 C; 4065 G; 4371 T; 0 U; 3 Other;
Query Match	
Best Local Similarity 20.2%; Score 1111; DB 3; Length 17111;	
Matches 1111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	936 A A A A G A A A A C G A A A T G A T A C A C C A A T C A G T C A A A A A A A G A T A T A T G G A G A C T A A G A C 995
DB	5745 A A A A G A A A A C G A A A T G A T A C A C C A A T C A G T C A A A A A A A G A T A T A T G G A G A C T A A G A C 5804
QY	996 G G T T C G T G T T C G T G C T G A C T T G C A C C A T A T C A T A A A A A T C G A A A C A G C A A A G A A T G C G C G 1055
DB	5805 G G T T C G T G T T C G T G C T G A C T T G C A C C A T A T C A T A A A A A T C G A A A C A G C A A A G A A T G C G C G 5864
QY	1056 A A A C G T A A A A G A A G T T A T G G A A A T A G A C T T A G A A G C A A A C T T A A G A G T G T T G T A T A G T 1115
DB	5865 A A A C G T A A A A G A A G T T A T G G A A A T A G A C T T A G A A G C A A A C T T A A G A G T G T T G T A T A G T 5924
QY	1116 G C A G T A T C T T A A A A T T T T G T A T A A T A G G A A T C G A A G T T T A A A T T A G A T T G C T F A A A A A T T T G T 1175
DB	5925 G C A G T A T C T T A A A A T T T T G T A T A A T A G G A A T T G A A G T T A A A T T A G A T T G C T A A A A A T T T G T 5984
QY	1176 A A T T A A G A A G A G A G T A T T A C A T G A A C A A A A A T A T A A A A T A T T C T A A A A C T T T T T A A C G A 1235
DB	5985 A A T T A A G A A G A G T A T T A C A T G A A C A A A A A T A T A A A T A T T C T C A A A C T T T T T A A C G A 6044
QY	1236 G T C A A A A A G T A C T C A C C A A A T A A T A A A C A A T T G A A T T T T A A A A G A A A C C G A T A C C G T T T 1295
DB	6045 G T G A A A A A G T A C T C A C C A A A T A A T A A A C A A T T G A A T T T T A A A A A A A C C G A T A C C G T T T 6104
QY	1296 A C G A A A T T G G A A C A G A G T A A A G G G C A T T T A C G A C G A A A C T G G C T A A A A T A A G T A A A C A G G 1355
DB	6105 A C G A A A T T G G A A C A G A G T A A A G G G C A T T T T A C G A C G A A A C T G G C T A A A A T A A G T A A A C A G G 6164
QY	1356 T A A C G T C T A T T G A A T T A G A C A G T C A C T A T T C A A C T T A T C G T C A G A A A A A T T A A A A C T G A 1415
DB	6165 T A A C G T C T A T T G A A T T A G A C A G T C A T C T A T T C A A C T T A T C G T C A G A A A A A T T A A A A C T G A 6224
QY	1416 A T A C T C G T G C A C T T T A A T T C A C A A G A T A T T C T A C A G T T T C A A T T C C C T T A A C A A A C A G A 1475
DB	6225 A T A C T C G T G T C A C T T T A A T T C A C A A G A T A T T C T A C A G T T T C A A T T C C C T T A A C A A C A G A 6284
QY	1476 G G T A T A A A A T G T G G A G T A T T C C T T A C A T T T A A G A C A C A A A T A T T A A A A A A G T G G 1535
DB	6285 G G T A T A A A A T G T T G G A G T A T T C C T T A C C A T T T A A G C A C A C A A A T A T T T A A A A A A G T G G 6344
QY	1536 T T T T T G A A A C C C A T G C G T C T G A C A T C T A T C T G A T T G T T G A A G A A G G A A T T C T A C A A G C G T A 1595
DB	6345 T T T T T G A A A C C C A T G C G T C T G A C A T C T A T C T G A T T G T T G A A G A A G G A T T C T A C A A G C G T A 6404
QY	1596 C C T T G G A T A T T C A C C G A A C A C T A G G T T G C T T T G C A C A C T C A A G T C T C A A T T C A G C A A T 1655
DB	6405 C C T T G G A T A T T C A C C G A A C A C T A G G T T G C T T T G C A C A C A C T C A A G T C T C A A T T C A G C A A T 6464
QY	1656 T G C T T A A G C T G C C A G C G G A A T G C T T T C A T C C T A A A C C A A A G T A A A C A G A T G T C T T A A T A A 1715
DB	6465 T G C T T A A G C T G C C A G C G G A A T G C T T T C A T C C T A A A C C A A A A G T A A A C A G A T G T C T T A A T A A 6524
QY	1716 A A C T T A C C C C C C A T A C C A C A G A T G T T C C A G A T A A A T T T G G A A G A C T A T A T A C G T A C T T T G 1775

5525	AACTTACCGCCCATACCCAGATGTTCCAGATAAAATATTGGAAGCTATATACGTACTTTG	6588		
1776	TTTCAAAATGGGTCAATCAGAAATATCGTCAACTGTTTACTAAAAATCAGTTTTCATCAAG	1835		
6585	TTTCAAAATGGGTCAATCAGAAATATCGTCAACTGTTTACTAAAAATCAGTTTTCATCAAG	6644		
1836	CAATGAACACACGCCAAAGTAAACAAATTTAAAGTACCGTTACTTTATGAGCAAGTATTGTCTA	1895		
6645	CAATGAACACACGCCAAAGTAAACAAATTTAAAGTACCGTTACTTTATGAGCAAGTATTGTCTA	6704		
1896	TTTTTAATAGTTTATCTATTATTTAAACGGGAGGAAATAATTTCTATGAGTCGCTTTTGTAAA	1955		
6705	TTTTTAATAGTTTATCTATTATTTAAACGGGAGGAAATAATTTCTATGAGTCGCTTTTGTAAA	6764		
1956	TTTGGAAAGTTACAGCTTACTTAAAGGGAATGATAGATAAAATTTATAGTATCTACTCTGACA	2015		
6765	TTTGGAAAGTTACAGCTTACTTAAAGGGAATGATAGATAAAATTTATAGTATCTACTCTGACA	6824		
2016	GCTTCCAAAGAGCTAAAGAGTCCCTAGCGC	2046		
6825	GCTTCCAAAGAGCTAAAGAGTCCCTAGCGC	6855		
RESULT 13				
AAD01290				
ID ID AAD01290 standard; DNA; 17116 BP.				
AAD01290;				
12-OCT-2000 (first entry)				
Arabidopsis transformation binary vector, pAct2Af-bin.				
Matrix Attachment Region; MAR; scaffold attachment region; pAct2Af-bin;				
Arabidopsis transformation binary vector; gene expression;				
transgenic organism; ds.				
Synthetic.				
Key	Location/Qualifiers			
misc_feature	64..709			
	/*tag= a			
	/note= "Corresponds to nucleotides 1-646 of MAR dimer-1"			
misc_feature	4044..4689			
	/*tag= b			
	/note= "Corresponds to nucleotides 1-646 of MAR dimer-3"			
WO200032800-A1.				
08-JUN-2000.				
30-NOV-1999; 99WO-US028123.				
01-DEC-1998; 98US-0110437P.				
(DOWC ) DOW AGROSCIENCES LLC.				
Van Der Geest AHM, Ainley WM, Cowen NM, Welter MB, Woosley AT;				
WPI; 2000-412345/35.				
An isolated DNA molecule for use as a matrix attachment region to				
increase expression of genes introduced in transformed plants comprises a				
298 base pair sequence described in the specification.				
Example 3; Page 64-68; 73pp; English.				
The patent discloses a DNA molecule, useful as matrix attachment region				
(MAR) or scaffold attachment region, to increase the expression of genes				
introduced in transformed plants. MARs are located in non-transcribed				
regions of genes and form the physical boundaries of individual DNA				
loops. They are rich in adenosine and thymine bases and contain certain				
conserved sequence elements and structural features. They are about 300-				

CC 2000 bp in length. Increased levels of expression of DNA introduced into  
 CC plants can be achieved by use of MAR. They can also reduce the position  
 CC effect in transgenic organisms. The present DNA sequence is the  
 CC Arabidopsis transformation binary vector pACT2af-bin, identical to the  
 CC vector pACT2-bin, except that it contains a MAR dimer-1 positioned 5' to  
 CC the Act2 transcription initiation region and the MAR dimer-3 positioned  
 CC 3' to the nos 3' UTR (untranslated region). This vector is used to test  
 CC the two orientations of the artificial MAR in Arabidopsis  
 XX  
 SQ Sequence 17116 BP; 4686 A; 3968 C; 4052 G; 4407 T; 0 U; 3 Other;  
 Query Match 20.2%; Score 1111; DB 3; Length 17116;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-212;  
 Matches 1111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 936 ARAAGAAACGAAATGATACACCATCAGTCGCAAAAGAGATATATCGGAGATAGAC 995  
 DB 5750 ARAAGAAACGAAATGATACACCATCAGTCGCAAAAGAGATATATCGGAGATAGAC 5809  
 QY 996 GGTTCGTGTTCTGCTGACATTCACCATATCATAAAATCGAAACAGCAAGAAATGGCGG 1055  
 DB 5810 GGTTCGTGTTCTGCTGACATTCACCATATCATAAAATCGAAACAGCAAGAAATGGCGG 5869  
 QY 1056 AAGCGTAAAGAGATGATGAAATAGACTTAGAAGCAAACTTAAGAGTGTGTATAGT 1115  
 DB 5870 AAGCGTAAAGAGATGATGAAATAGACTTAGAAGCAAACTTAAGAGTGTGTATAGT 5929  
 QY 1116 GCAGTATCTTAAATTTGATATAGGAATTCAGTTAAATAGATGCTAAAAATTTGT 1175  
 DB 5930 GCAGTATCTTAAATTTGATATAGGAATTCAGTTAAATAGATGCTAAAAATTTGT 5989  
 QY 1176 AATTAAGAGGAGTATCATGAACAAATAATAAATATCTCAAACTTTTAAACGA 1235  
 DB 5990 AATTAAGAGGAGTATCATGAACAAATAATAAATATCTCAAACTTTTAAACGA 6049  
 QY 1236 GTGAAAGTACTCAACCAATATTAACCAATTTGATTTAAAGAAACCGATACCGTTT 1295  
 DB 6050 GTGAAAGTACTCAACCAATATTAACCAATTTGATTTAAAGAAACCGATACCGTTT 6109  
 QY 1296 ACGAAATTTGAACAGGTAAAGGGGCAATTTAAGCAAGAACTGGCTAAATAAGTAAACAGG 1355  
 DB 6110 ACGAAATTTGAACAGGTAAAGGGGCAATTTAAGCAAGAACTGGCTAAATAAGTAAACAGG 6169  
 QY 1356 TAAGCTGATTTGAATAGACAGTCACTATCTCACTATGCTGAGAAATTTAAACTGA 1415  
 DB 6170 TAAGCTGATTTGAATAGACAGTCACTATCTCACTATGCTGAGAAATTTAAACTGA 6229  
 QY 1416 ATACTGCTGTCACCTTAAATTTCAACAGATATTTCAAGTTTCAATTCCTTAACAAACAGA 1475  
 DB 6230 ATACTGCTGTCACCTTAAATTTCAACAGATATTTCAAGTTTCAATTCCTTAACAAACAGA 6289  
 QY 1476 GGTATAAAATTTGGGAGTATTCCTTACCAATTTAAGCAACAAATTTAAAAAAGTGG 1535  
 DB 6290 GGTATAAAATTTGGGAGTATTCCTTACCAATTTAAGCAACAAATTTAAAAAAGTGG 6349  
 QY 1536 TTTTGAAGCCATCGCTGTCATCTATCTGATTTGTTGAAGAGATTTCAACAGGTA 1595  
 DB 6350 TTTTGAAGCCATCGCTGTCATCTATCTGATTTGTTGAAGAGATTTCAACAGGTA 6409  
 QY 1596 CTTTGGATTTTCCCGAACACTAGGTTGCTTCTTGCACTCAAGTCTCGATTCAAGAT 1655  
 DB 6410 CTTTGGATTTTCCCGAACACTAGGTTGCTTCTTGCACTCAAGTCTCGATTCAAGAT 6469  
 QY 1656 TGCTTAAGCTGCGACGGGATGCTTTCATCTTAACCAAGTAAACAGTGTCTTAATA 1715  
 DB 6470 TGCTTAAGCTGCGACGGGATGCTTTCATCTTAACCAAGTAAACAGTGTCTTAATA 6529  
 QY 1716 AACTTACCGCCATPACCAAGATGTTCCAGATAAATTTGGAAGCTATATACGTACTTTG 1775  
 DB 6530 AACTTACCGCCATPACCAAGATGTTCCAGATAAATTTGGAAGCTATATACGTACTTTG 6589  
 QY 1776 TTTCAAAATGGTCAATCGAGATATCGTCAACTGTTTACTAAAAATCAAGTTTCATCAAG 1835

DB 6590 TTTCAAAATGGTCAATCGAGATATCGTCAACTGTTTACTAAAAATCAGTTTCATCAAG 6649  
 QY 1836 CAATGAAACACGCCAAAGTAAACAAATTTAAGTACCGTTACTATGAGCAAGTATTGCTA 1895  
 DB 6650 CAATGAAACACGCCAAAGTAAACAAATTTAAGTACCGTTACTATGAGCAAGTATTGCTA 6709  
 QY 1896 TTTTAAATAGTATTCTATTTAATTTAAGCGGAGGAAATATTTCTATGAGTCGCTTTGTA 1955  
 DB 6710 TTTTAAATAGTATTCTATTTAATTTAAGCGGAGGAAATATTTCTATGAGTCGCTTTGTA 6769  
 QY 1956 TTTGAAAGTTACAGTTTACTTAAAGGGAATAGATAAAATTTATTAGTATCTACTGACA 2015  
 DB 6770 TTTGAAAGTTTACAGTTTACTTAAAGGGAATAGATAAAATTTATTAGTATCTACTGACA 6829  
 QY 2016 GTTCCCAAGGAGCTAAAGAGGTCCTTAGCGC 2046  
 DB 6830 GTTCCCAAGGAGCTAAAGAGGTCCTTAGCGC 6860  
 RESULT 14  
 ID ABK98591/c  
 ID ABK98591 standard; DNA; 10929 BP.  
 AC ABK98591;  
 XX 07-AUG-2003 (revised)  
 DT 21-OCT-2002 (first entry)  
 DE pEP25 vector fragment containing Xylr/Xylo/CP25 sequences.  
 XX ds; promoter; gram positive bacteria; fusion promoter; T5; CP25; P32;  
 KW P59; P1P2; PL; xylO; teco; trpO; malo; lambda; lambda; lambda; lambda;  
 KW antibiotic; vector.  
 XX Eubacteria.  
 OS Bacteriophage lambda.  
 OS Escherichia coli.  
 OS Synthetic.  
 XX WO200251982-A2.  
 PN 04-JUL-2002.  
 XX 21-DEC-2001; 2001WO-US050250.  
 PF 27-DEC-2000; 2000US-0259434P.  
 PR 06-SEP-2001; 2001US-00948993.  
 XX (ELIT-) ELITRA PHARM INC.  
 PA Haselbeck R, Wall D, Gross M;  
 PI WPI; 2002-575374/61.  
 DR Isolated nucleic acid comprises bacterial promoters modified to have  
 XX altered activity in at least one gram-positive organism, e.g. Bacillus  
 PT antracis or Clostridium botulinum, useful for regulating gene expression  
 PT in bacteria.  
 XX Example 1; Page 202-205; 246pp; English.  
 PS The invention relates to an isolated nucleic acid comprising a fusion  
 XX promoter comprising at least one promoter that is modified to have  
 CC altered activity in at least one gram-positive organism, or comprising  
 CC T5, CP25, P32, P59, P1P2 or PL linked to at least one operator consisting  
 CC of xylO, teco, trpO, malo or lambda; lambda; lambda; lambda; lambda;  
 CC positioned so binding of a repressor to an operator represses  
 CC host cells comprising the fusion promoter. Also included are vectors and  
 CC involved in cellular proliferation or required for proliferation of a  
 CC prokaryotic cell using the vector, a method of identifying compounds that  
 CC inhibit the proliferation of a prokaryotic cell using the vector, a  
 CC method of identifying a compound that reduces the activity or level of a

CC	gene product required for proliferation of a cell using the vector, a
CC	compound identified by the methods, a method of inhibiting the activity
CC	or expression of a gene in an operon required for proliferation using the
CC	vector, manufacturing an antibiotic comprising using the vector or cell.
CC	and identifying a nucleic acid with promoter activity in <i>Enterococcus</i>
CC	<i>faecalis</i> . The fusion promoters are useful for regulating nucleic acid or
CC	polypeptide expression, particularly for regulating gene expression in
CC	bacteria and for identifying proliferation-required genes or molecules
CC	with potential antibiotic activity. The modified promoters are also
CC	useful for replacing endogenous promoters to create cells with specific
CC	regulatable genes. The present sequence is vector (or fragment)
CC	incorporating a fusion promoter sequence of the invention. (Updated on 07
CC	-AUG-2003 to correct OS field.)
XX	
SQ	Sequence 10929 BP; 3259 A; 2175 C; 2157 G; 3338 T; 0 U; 0 Other;
	Query Match 20.2%; Score 1109.4; DB 6; Length 10929;
	Best Local Similarity 99.9%; Pred. No. 6e-212;
	Matches 1110; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY	936 AAAAGAAAAAGAAATGATACACCAATCAGTCGAAAAAAGATATAATGGGAGATAAGAC 995
DB	5830 AAAAGAAAAAGAAATGATACACCAATCAGTCGAAAAAAGATATAATGGGAGATAAGAC 5771
QY	996 GGTTCGTGTCGTGCTGACTTGCACCATATCATAAAAATCGAAACAGCAAGAAATGGCGG 10555
DB	5770 GGTTCGTGTCGTGCTGACTTGCACCATATCATAAAAATCGAAACAGCAAGAAATGGCGG 5711
QY	1056 AAAAGTAAAAAGAGTTATGGAANAATAGACCTTAGAGCAAACTTAGAGTGTGTCATAGT 1115
DB	5710 AAAAGTAAAAAGAGTTATGGAANAATAGACCTTAGAGCAAACTTAGAGTGTGTCATAGT 5651
QY	1116 GCAGTATCTTAAAAATTTGTATATATAGGAATTCGAAGTTAAATTTAGATGCTAAAAATTGTT 1175
DB	5650 GCAGTATCTTAAAAATTTGTATATATAGGAATTCGAAGTTAAATTTAGATGCTAAAAATTGTT 5591
QY	1176 AATTAAAGAGGAGTGATTACATGAACAAAAATATAAAATATTCTCAAAACTTTTTTAACGA 1235
DB	5590 AATTAAAGAGGAGTGATTACATGAACAAAAATATAAAATATTCTCAAAACTTTTTTAACGA 5531
QY	1236 GTGAAAAAGTACTCAACCAATATATAACCAATTTGAATTTTAAAGNAAACCGATACCGTTT 1295
DB	5530 GTGAAAAAGTACTCAACCAATATATAACCAATTTGAATTTTAAAGNAAACCGATACCGTTT 5471
QY	1296 ACGAAATTTGGAACAGGTAAAGGGCAATTTAACAGCGAAACTGGCTAAAAATAGTAAACAGG 1355
DB	5470 ACGAAATTTGGAACAGGTAAAGGGCAATTTAACAGCGAAACTGGCTAAAAATAGTAAACAGG 5411
QY	1356 TTAACGTCTATTGAATTAGACAGTCATCTTAACTTATCGTCAGAAAAATTTAAACCTGA 1415
DB	5410 TTAACGTCTATTGAATTTAGACAGTCATCTTAACTTATCGTCAGAAAAATTTAAACCTGA 5351
QY	1416 ATACTCGTGTGCATTTTAATTCACCAAGATATCTCAGTTTCAATTTCCCTTAACAAACAGA 1475
DB	5350 ATACTCGTGTGCATTTTAATTCACCAAGATATCTCAGTTTCAATTTCCCTTAACAAACAGA 5291
QY	1476 GGTATAAAATTTGTGGGAGTATTCCTTCAATTTAAAGCAACAATATTATTAAGAGTGG 1535
DB	5290 GGTATAAAATTTGTGGGAGTATTCCTTCAATTTAAAGCAACAATATTATTAAGAGTGG 5231
QY	1536 TTTTGTGAAGCCATGCGTCTGACATCTATCTGATTTGTGAAGAAGGATTTCTACAAGCGTA 1595
DB	5230 TTTTGTGAAGCCATGCGTCTGACATCTATCTGATTTGTGAAGAAGGATTTCTACAAGCGTA 5171
QY	1596 CCTTCGGATATTCAACGAACACTAGGGTGTCTTTGACACACTCAAGTCTCGATTCAGCAAT 1655
DB	5170 CCTTCGGATATTCAACGAACACTAGGGTGTCTTTGACACACTCAAGTCTCGATTCAGCAAT 5111
QY	1656 TCGTTAAAGTGGCAGCGGAATGCTTTCATCCTTAAACCAAAAGTAAACAGTGTCTTAATAA 1715
DB	5110 TCGTTAAAGTGGCAGCGGAATGCTTTCATCCTTAAACCAAAAGTAAACAGTGTCTTAATAA 5051
QY	1716 AACTTTACCGCGCATACACAGATGCTCCAGATATAATTTGGAAGCTATATACGTACTTTG 1775

Db	5050	AACTTACCCGCCATACACAGATGTTCCAGATAAATATTGGAAGCTATATACGTACTTTG	4991
Qy	1776	TTTCAAAATGGGTCAATCGAGAAATATCGTCAACTGTTTACTAAAAATCAAGTTTCATCAAG	1835
Db	4990	TTTCAAAATGGGTCAATCGAGAAATATCGTCAACTGTTTACTAAAAATCAAGTTTCATCAAG	4931
Qy	1836	CAATGAACACGCCAAAGTAACAAATTAAGTACCGTTACTTATGAGCAAGTATTGCTTA	1895
Db	4930	CAATGAACACGCCAAAGTAACAAATTAAGTACCGTTACTTATGAGCAAGTATTGCTTA	4871
Qy	1896	TTTTTAATAGTTATCTATTATTATTAACGGGAGGAATAATTCATGAGTCGCTTTTGAAA	1955
Db	4870	TTTTTAATAGTTATCTATTATTATTAACGGGAGGAATAATTCATGAGTCGCTTTTGAAA	4811
Qy	1956	TTTGAAAGTTACACGTTTACTTAAAGGAATGTAGATAAATATTAGGTATCTACTGACA	2015
Db	4810	TTTGAAAGTTACACGTTTACTTAAAGGAATGTAGATAAATATTAGGTATCTACTGACA	4751
Qy	2016	GCITTCAGGAGCTAAAGAGGTCCTTAGCGC	2046
Db	4750	GCITTCAGGAGCTAAAGAGGTCCTTAGCGC	4720
RESULT 15			
ACD13842/c			
ID	ACD13842	standard; DNA; 10929 BP.	
AC	ACD13842;		
XX			
DT	15-AUG-2003	(first entry)	
XX			
DE	Plasmid pEP25	containing L. lactis CP25 promoter.	
XX			
KW	Promoter; ds;	gram positive bacteria; Staphylococcus aureus; plasmid;	
KW	Enterococcus faecalis;	operator; xylO; tetO; trpO; malO; lambda-clo;	
KW	cellular proliferation.		
XX			
OS	Lactococcus lactis.		
OS	Synthetic.		
XX			
PN	US2003027286-A1.		
XX			
PD	06-FEB-2003.		
XX			
PF	21-DEC-2001; 2001US-00032393.		
XX			
PR	06-SEP-2000; 2000US-0230335P.		
PR	27-DEC-2000; 2000US-0259434P.		
XX			
PA	(HASE// HASELBECK R.		
PA	(WALL// WALL D.		
PA	(GROS// GROSS M.		
PI	Haselbeck R, Wall D, Gross M;		
XX			
DR	WPI; 2003-479541/45.		
XX			
PT	New isolated nucleic acid	comprising a fusion promoter having at least	one promoter that is modified to have altered activity in at least one gram-positive organism, useful for regulating gene expression in bacteria.
PT			
PT			
PT			
XX			
FS	Example 1; Page 58-62; 142pp; English.		
XX			
CC	The invention relates to an isolated nucleic acid	comprising a fusion promoter having at least one promoter that is modified to have altered activity in at least one gram-positive organism (e.g. Staphylococcus aureus or Enterococcus faecalis). The promoter is linked to at least one operator selected from xylO, tetO, trpO, malO and lambda-clo, which are positioned such that the binding of at least one repressor to the operator represses transcription from the fusion promoter. Also included are a vector comprising the isolated nucleic acid, a host cell comprising	

the nucleic acid. The fusion promoter is useful for identifying genes involved in cellular proliferation, identifying a compound that reduces the activity or level of a gene product required for proliferation of a cell, inhibiting the activity or expression of a gene in an operon required for proliferation, manufacturing an antibiotic, identifying a gene that is required for proliferation of a prokaryotic cell, identifying a compound that inhibits the proliferation of a prokaryotic cell and regulating gene expression in bacteria. The present sequence is a plasmid containing a fusion promoter of the invention

XX  
SQ Sequence 10929 BP; 3259 A; 2175 C; 2157 G; 3338 T; 0 U; 0 Other;

Query Match 20.2%; Score 1109.4; DB 8; Length 10929;  
Best Local Similarity 99.9%; Pred. No. 6e-212;  
Matches 1110; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 936 AAAAGAAAAACGAAATGATACCAATCAGTCGCAAAAAAGATATTAATGGAGATAAGAC 995  
DB 5830 AAAAGAAAAACGAAATGATACCAATCAGTCGCAAAAAAGATATTAATGGAGATAAGAC 5771

QY 996 GGTTCGTGTTGCTGCTGACTGACCATATCATAAATCGAACACGCAAGATCGCG 1055  
DB 5770 GGTTCGTGTTGCTGCTGACTGACCATATCATAAATCGAACACGCAAGATCGCG 5711

QY 1056 AAACGTAAAAAGATTATGAAATAGACTTAGAACCAACTTAAAGATGTTGTATAGT 1115  
DB 5710 AAACGTAAAAAGATTATGAAATAGACTTAGAACCAACTTAAAGATGTTGTATAGT 5651

QY 1116 GCAGTATCTTAAATTTTGTATATAGGAATGAAGTTAAATAGATGCTAAAAATTTGT 1175  
DB 5650 GCAGTATCTTAAATTTTGTATATAGGAATGAAGTTAAATAGATGCTAAAAATTTGT 5591

QY 1176 AATTAAGAGAGATGATTACATGAACAAAAATATAAATTTCTCAAACTTTTAAACGA 1235  
DB 5590 AATTAAGAGAGATGATTACATGAACAAAAATATAAATTTCTCAAACTTTTAAACGA 5531

QY 1236 GTGAAAAAGTACTCAACCAATATAAACAATTTAAAGAAACCGATACCGTTT 1295  
DB 5530 GTGAAAAAGTACTCAACCAATATAAACAATTTAAAGAAACCGATACCGTTT 5471

QY 1296 ACCGAATTTGGAACAGGTAAAGGCTTTTAAACGACGAACTGGCTAAATTAAGTAACAG 1355  
DB 5470 ACCGAATTTGGAACAGGTAAAGGCTTTTAAACGACGAACTGGCTAAATTAAGTAACAG 5411

QY 1356 TAACGTCTATTGAATTAGACAGTCATCTATTCAACTTATCGTCAGAAAAATTTAAACTGA 1415  
DB 5410 TAACGTCTATTGAATTAGACAGTCATCTATTCAACTTATCGTCAGAAAAATTTAAACTGA 5351

QY 1416 ATACTCGTGTCACTTTAATTCACCAAGATATTTCAAGTTTCAATTCCTTCAACACAGA 1475  
DB 5350 ATACTCGTGTCACTTTAATTCACCAAGATATTTCAAGTTTCAATTCCTTCAACACAGA 5291

QY 1476 GGTATAAAATTTGTTGGAGTATTCCTTACCATTAAAGCACACAAATTTAAAGAAAGTGG 1535  
DB 5290 GGTATAAAATTTGTTGGAGTATTCCTTACCATTAAAGCACACAAATTTAAAGAAAGTGG 5231

QY 1536 TTTTGTAAAGCCATGCGCTGACATCTATCTGATTTGTTGAAGAGATTTCTCAAGCGTA 1595  
DB 5230 TTTTGTAAAGCCATGCGCTGACATCTATCTGATTTGTTGAAGAGATTTCTCAAGCGTA 5171

QY 1596 CTTTGGATATTTCACGGAACACTAGGTTGCTTTGACACCTCAAGTCTCGATTTCGCAAT 1655  
DB 5170 CTTTGGATATTTCACGGAACACTAGGTTGCTTTGACACCTCAAGTCTCGATTTCGCAAT 5111

QY 1656 TGCTTAAGCTGCCAGCGGAATGCTTTTCATCTAAACCAAAAGTAAACAGTGTCTTAATAA 1715  
DB 5110 TGCTTAAGCTGCCAGCGGAATGCTTTTCATCTAAACCAAAAGTAAACAGTGTCTTAATAA 5051

QY 1716 AACTTACCCGCCATACACAGATGTTCCAGATATAATTTGGAAGCTATATACGACTTTG 1775  
DB 5050 AACTTACCCGCCATACACAGATGTTCCAGATATAATTTGGAAGCTATATACGACTTTG 4991

QY 1776 TTTCAAAATGGGTCAATTCGAGAAATATCGTCAACTGTTTACTAAAAATCAGTTTTCATCAAG 1835

Search completed: September 1, 2004, 15:00:05  
Job time : 1900 secs

DB 4990 TTTCAAAATGGGTCAATCGAATATCGTCACTGTTTACTAAAAATCAGTTTTCATCAAG 4931  
QY 1836 CAATGAAACACGCCAAAGTAAACAAATTTAAAGTACCGTTTACTTATGAGCAAGTATTGTCTA 1895  
DB 4930 CAATGAAACACGCCAAAGTAAACAAATTTAAAGTACCGTTTACTTATGAGCAAGTATTGTCTA 4871  
QY 1896 TTTTAAATAGTTATCTATTTTAAACGGGAGGAAATAATTTCTATGAGTCGCTTTTGTAAA 1955  
DB 4870 TTTTAAATAGTTATCTATTTTAAACGGGAGGAAATAATTTCTATGAGTCGCTTTTGTAAA 4811  
QY 1956 TTTGAAAAATTACAGCTTACTTAAAGGGAATGTAGATAAATTTATTAGGTATACTACTGACA 2015  
DB 4810 TTTGAAAAATTACAGCTTACTTAAAGGGAATGTAGATAAATTTATTAGGTATACTACTGACA 4751  
QY 2016 GCTTCCAAAGGAGCTAAAGAGTCCCTAGCGC 2046  
DB 4750 GCTTCCAAAGGAGCTAAAGAGTCCCTAGCGC 4720

This Page Blank (uspto)



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 1, 2004, 14:22:19 ; Search time 336 Seconds

(without alignments)  
9079.063 Million cell updates/sec

Title: US-10-030-390-2

Perfect score: 5497

Sequence: 1 gaattcgattaaagtcattt.....ctagtcacaaagtattataata 5497

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: Issued Patents NA.\*
- 2: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*
- 1: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*
- 3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*
- 4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*
- 5: /cgn2\_6/ptodata/2/ina/PTUS\_COMB.seq.\*
- 6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4994.4	90.9	5217	4	US-09-100-703A-25
2	4993.4	90.8	5230	4	US-09-100-703A-26
3	4944.6	90.0	5231	4	US-09-100-703A-27
4	1111	20.2	10323	4	US-09-280-428A-11
5	884	16.1	6243	2	US-08-056-075-1
6	737.2	13.4	783	4	US-09-134-000C-3303
7	731.6	13.3	738	2	US-08-743-637B-193
8	273.2	5.0	1167	4	US-09-134-000C-3296
9	247.2	4.5	696	4	US-09-134-000C-3315
10	244.2	4.4	3792	2	US-08-992-334-1
11	244.2	4.4	3792	3	US-08-302-752-1
12	244.2	4.4	5234	2	US-08-992-334-2
13	244.2	4.4	5234	3	US-08-302-752-2
14	244.2	4.4	5313	4	US-09-329-920-1
15	244.2	4.4	6169	2	US-08-875-154-2
16	244.2	4.4	6722	2	US-08-992-334-3
17	244.2	4.4	6722	3	US-08-302-752-3
18	244.2	4.4	10216	2	US-08-875-154-1
19	241	4.4	735	2	US-08-743-637B-194
20	240.8	4.4	720	4	US-09-107-532A-982
21	229.2	4.2	732	2	US-08-743-637B-192
22	223.2	4.1	2964	3	US-08-286-819A-18
23	223.2	4.1	2964	3	US-08-980-357-18
24	223.2	4.1	3190	2	US-08-286-819A-30
25	223.2	4.1	3190	3	US-08-980-357-30
26	223.2	4.1	10851	2	US-08-826-819A-16
27	223.2	4.1	10851	3	US-08-980-357-16

28	191	3.3	204	4	US-09-134-000C-3302	Sequence 3302, App
29	143.2	2.6	384	1	US-08-040-753-1	Sequence 1, Appli
30	143.2	2.6	5502	4	US-09-702-705-785	Sequence 785, App
31	143.2	2.6	5502	4	US-09-736-457-785	Sequence 785, App
32	143.2	2.6	5502	4	US-09-614-124B-785	Sequence 785, App
33	143.2	2.6	5502	4	US-09-671-325-785	Sequence 785, App
34	143.2	2.6	5502	4	US-09-589-184-785	Sequence 785, App
35	142	2.6	142	4	US-09-060-878-1	Sequence 1, Appli
36	141.6	2.6	2220	2	US-08-864-224-1	Sequence 1, Appli
37	140.8	2.6	4557	4	US-08-778-717-5	Sequence 5, Appli
38	140.8	2.6	6363	2	US-08-929-967-6	Sequence 6, Appli
39	140	2.5	7659	3	US-09-128-314-4	Sequence 4, Appli
40	140	2.5	8157	3	US-09-128-314-3	Sequence 3, Appli
C 41	139.6	2.5	6361	3	US-08-646-538-7	Sequence 7, Appli
C 42	139.6	2.5	6361	3	US-09-503-223-7	Sequence 7, Appli
43	139.6	2.5	6727	2	US-08-125-462-2	Sequence 2, Appli
44	139.6	2.5	6727	2	US-08-891-848-2	Sequence 2, Appli
45	139.6	2.5	6799	2	US-08-125-462-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1  
US-09-100-703A-25  
; Sequence 25, Application US/09100703A  
; Patent No. 6610300  
; GENERAL INFORMATION:  
; APPLICANT: SEGERS, Ruud PAM  
; APPLICANT: WATERFIELD, Nicolas R  
; APPLICANT: FRANDSEN, Peer L  
; APPLICANT: WELLS, Jeremy M.  
; TITLE OF INVENTION: COLISTRIDIMUM PERFRINGENS VACCINE  
; FILE REFERENCE: 97288 US  
; CURRENT APPLICATION NUMBER: US/09/100,703A  
; CURRENT FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: EP97201888.1  
; PRIOR FILING DATE: 1997-06-20  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 25  
; LENGTH: 5217  
; TYPE: DNA  
; ORGANISM: Clostridium perfringens  
US-09-100-703A-25

Query Match	90.9%	Score	4994.4	DB	4	Length	5217
Best Local Similarity	99.9%	Pred. No.	0	Mismatches	6	Indels	0
Matches	4998	Conservative	0	Matches	6	Indels	0
QY	494	TTCTAACTAGTAGTCCGGCTGCTTAAAGCCGCGAAGGAGCTGAGTTGGCTGCTGCC	553				
DB	214	TCCAAATGATCAGATCCGGCTGCTTAAAGCCGCGAAGGAGCTGAGTTGGCTGCTGCC	273				
QY	554	ACCGCTGAGCAATACTAGCATAACCCCTTGGGGCTCTAAACGGTCTTCGAGGGGTTTT	613				
DB	274	ACCGCTGAGCAATACTAGCATAACCCCTTGGGGCTCTAAACGGTCTTCGAGGGGTTTT	333				
QY	614	TTGCTGAAAGGAGGAACCTATATCCGGATGACCTCGAGCAAGCTCTAGAAATCGATACGAT	673				
DB	334	TTGCTGAAAGGAGGAACCTATATCCGGATGACCTCGAGCAAGCTCTAGAAATCGATACGAT	393				
QY	674	TTTGAAGTGGCAACAGATAAAAAAGCAGTTTAAATTTGTTGCTGACTTTTAAACAA	733				
DB	394	TTTGAAGTGGCAACAGATAAAAAAGCAGTTTAAATTTGTTGCTGACTTTTAAACAA	453				
QY	734	GCAATACAAATCATTGTCGCAACAGATAGCGACAGAGAGCGGAAAAACATTGCTGGTCG	793				
DB	454	GCAATACAAATCATTGTCGCAACAGATAGCGACAGAGAGCGGAAAAACATTGCTGGTCG	513				
QY	794	ATCATTGATAAGCAATGCTTTTCTAAAGATAAAGCTATAAAGACTATGATCAT	853				
DB	514	ATCATTGATAAGCAATGCTTTTCTAAAGATAAAGCTATAAAGACTATGATCAT	573				



3814	Db	TTGCTCGCATACAAGAACGGACAAATGTAGAAATTTTTTGATCCTAATTAACGGTTATTCCTT	3873
4154	QY	TCAGAAATTCGCAAGATTGGTCTCTTTCAAAACAAACAGATAATTAAGGGCTTTTACTCGTTCAA	4213
3874	Db	TCAGAAATTCGCAAGATTGGTCTCTTTCAAAACAAACAGATAATTAAGGGCTTTTACTCGTTCAA	3933
4214	QY	GTCCTACGGTTTTTAACCGGTACAGAAAGGCAAAAAACAAGTAGATGAACCGTGGTTAAATC	4273
3934	Db	GTCCTACGGTTTTTAACCGGTACAGAAAGGCAAAAAACAAGTAGATGAACCGTGGTTAAATC	3993
4274	QY	TCTTATTGCAGAAACGAAATTTTCAGGAGAAAAGGGTTTTAATAGGGCGGTAAATAACGCTCA	4333
3994	Db	TCTTATTGCAGAAACGAAATTTTCAGGAGAAAAGGGTTTTAATAGGGCGGTAAATAACGCTCA	4053
4334	QY	TGTTTACCTCTCTTTAGGCTACTTTAGTTTCAGGCTATTCAATTCGAAACGTCGGAATATA	4393
4054	Db	TGTTTACCTCTCTTTAGGCTACTTTAGTTTCAGGCTATTCAATTCGAAACGTCGGAATATA	4113
4394	QY	ATATGTTTCAGTTTTAATAATCGATTAGATCAACCCCTTGAAGAGAAAAGAGTAATCAAAA	4453
4114	Db	ATATGTTTCAGTTTTAATAATCGATTAGATCAACCCCTTGAAGAGAAAAGAGTAATCAAAA	4173
4454	QY	TTGTTAGAAAGTCCTATTTCAGAAAACTATCAAGGGCTTAATAGGGAATACATTACCATTC	4513
4174	Db	TTGTTAGAAAGTCCTATTTCAGAAAACTATCAAGGGCTTAATAGGGAATACATTACCATTC	4233
4514	QY	TTTGCAAAAGCTTTGGGTATCAAGTGATTTAAACAGATAAGATTATTTTGTCCGCTCAAGGCT	4573
4234	Db	TTTGCAAAAGCTTTGGGTATCAAGTGATTTAAACAGATAAGATTATTTTGTCCGCTCAAGGCT	4293
4574	QY	GGTTTTAAATTCAGAAAAAAGAACGCAACGCTCAACGCTTCATTGTGCAGAAATGGAAG	4633
4294	Db	GGTTTTAAATTCAGAAAAAAGAACGCAACGCTCAACGCTTCATTGTGCAGAAATGGAAG	4353
4634	QY	AGAGTTTAATGCGTTATATTACGGAAAAAAGCGATGTATACAAGCCTTATTTAGTGACGA	4693
4354	Db	AGAGTTTAATGCGTTATATTACGGAAAAAAGCGATGTATACAAGCCTTATTTAGTGACGA	4413
4694	QY	CCAAAAACAGATTAGAGAAAGTCTAGGCAATTCCTGAACGGACATTTAGATAAATTCGTCGA	4753
4414	Db	CCAAAAACAGATTAGAGAAAGTCTAGGCAATTCCTGAACGGACATTTAGATAAATTCGTCGA	4473
4754	QY	AGGTACTGAAGCGAATCAGGAAATTTCTTTTAAGATTAAACCGAGAGAAATTCGTGGCA	4813
4474	Db	AGGTACTGAAGCGAATCAGGAAATTTCTTTTAAGATTAAACCGAGAGAAATTCGTGGCA	4533
4814	QY	TTCAACTGCTAGTGTATAATCATTTGCTATCCGATCATTAAGATGAATAAAGAGAGAAA	4873
4534	Db	TTCAACTGCTAGTGTATAATCATTTGCTATCCGATCATTAAGATGAATAAAGAGAGAAA	4593
4874	QY	AAGAAAGCTATATAAAGGCGCTGACAAATTCCTTTTGACTTAGAGCATACATTCATTCAAG	4933
4594	Db	AAGAAAGCTATATAAAGGCGCTGACAAATTCCTTTTGACTTAGAGCATACATTCATTCAAG	4653
4934	QY	AGACITTTAAACAAGCTACGAGAACGCCCTTAAACGGACACACAACCTCGATTTGTTTAGCT	4993
4654	Db	AGACITTTAAACAAGCTACGAGAACGCCCTTAAACGGACACACAACCTCGATTTGTTTAGCT	4713
4994	QY	ATGATACAGGCTGAAAAATAAAACCCGCCACTATGCCATTATCTATCTATGATACGCTG	5053
4714	Db	ATGATACAGGCTGAAAAATAAAACCCGCCACTATGCCATTATCTATCTATGATACGCTG	4773
5054	QY	TTTGTTTTTCTTTGCTGTTTAGCGAATGATTAGCAGAAAAATATACAGAGTAGAGTTTAA	5113
4774	Db	TTTGTTTTTCTTTGCTGTTTAGCGAATGATTAGCAGAAAAATATACAGAGTAGAGTTTAA	4833
5114	QY	TTTAATTTTAGGGGAGAGGAGAGTAGCCCGAAAACCTTTTAGTTGGCTTGGACTGAA	5173
4834	Db	TTTAATTTTAGGGGAGAGGAGAGTAGCCCGAAAACCTTTTAGTTGGCTTGGACTGAA	4893
5174	QY	CGAAGTGAGGAAAAGGCTACTTAAACGTCGAGGGCGAGTGAGAGCGAAGCGCAACACTTGA	5233
4894	Db	CGAAGTGAGGAAAAGGCTACTTAAACGTCGAGGGCGAGTGAGAGCGAAGCGCAACACTTGA	4953

QY 5234 TTTTAAATTTCTATCTTTATAGTCAATAGAGTATATCTTATTGCTCCTATAACTAT 5293  
DB 4954 TTTTAAATTTCTATCTTTATAGTCAATAGAGTATATCTTATTGCTCCTATAACTAT 5013  
QY 5294 TTAGCAGCATATAGATTTATTGAATAGGTCATTAAAGTTGAGCATATTAGAGGAGAAA 5353  
DB 5014 TTAGCAGCATATAGATTTATTGAATAGGTCATTAAAGTTGAGCATATTAGAGGAGAAA 5073  
QY 5354 ATCTTGAGAAATATTGAAGAACCCGATTAATGAGTTGAGTTAGTTCTTGGTTACG 5413  
DB 5074 ATCTTGAGAAATATTGAAGAACCCGATTAATGAGTTGAGTTAGTTCTTGGTTACG 5133  
QY 5414 TGGTTTTTAACATAAAGTAGTGAATTTTGGTTTGGTTGTTGTTGTTGTTAGTA 5473  
DB 5134 TGGTTTTTAACATAAAGTAGTGAATTTTGGTTTGGTTGTTGTTGTTGTTAGTA 5193  
QY 5474 TTTGCTAGTCAAGTGAATTAATA 5497  
DB 5194 TTTGCTAGTCAAGTGAATTAATA 5217

RESULT 2  
US-09-100-703A-26  
; Sequence 26, Application US/09100703A  
; Patent No. 6610300  
; GENERAL INFORMATION:  
; APPLICANT: SEGERS, Ruud PAM  
; APPLICANT: WATERFIELD, Nicolas R  
; APPLICANT: FRANDSEN, Peer L  
; APPLICANT: WELLS, Jeremy M.  
; TITLE OF INVENTION: COLISTRIDIUM PERFRINGENS VACCINE  
; FILE REFERENCE: 97288 US  
; CURRENT APPLICATION NUMBER: US/09/100,703A  
; CURRENT FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: EP97201888.1  
; PRIOR FILING DATE: 1997-06-20  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 26  
; TYPE: DNA  
; ORGANISM: Clostridium perfringens  
US-09-100-703A-26

Query Match 90.8%; Score 4993.4; DB 4; Length 5230;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 4997; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 495 TCTAACTAGTAGTCGGCTCTCAACAAAGCCGAAAGGAGCTGAGTTGGCTCTGCCA 554  
DB 228 TCTTAGAATGATCCGGCTCTCAACAAAGCCGAAAGGAGCTGAGTTGGCTCTGCCA 287  
QY 555 CGCTGAGCAATACTAGCATAACCCCTTGGGGCTCTAAACGGGTCTTGAGGGGTTTT 614  
DB 288 CGCTGAGCAATACTAGCATAACCCCTTGGGGCTCTAAACGGGTCTTGAGGGGTTTT 347  
QY 615 TGCTGAAGGAGGAGTATATCCGGATGACCTGCGAGGAGCTCTAGAAATGATACGATT 674  
DB 348 TGCTGAAGGAGGAGTATATCCGGATGACCTGCGAGGAGCTCTAGAAATGATACGATT 407  
QY 675 TTGAAGTGGCAACAGATATAAAGAGCAGTTTAAATTTGTTGCTGCAACTTTTAAACAAG 734  
DB 408 TTGAAGTGGCAACAGATATAAAGAGCAGTTTAAATTTGTTGCTGCAACTTTTAAACAAG 467  
QY 735 CAAATACATATCTTGGCAACAGATAGCAGAGAGGCGAAACATTTGCTTGGTGA 794  
DB 468 CAAATACATATCTTGGCAACAGATAGCAGAGAGGCGAAACATTTGCTTGGTGA 527  
QY 795 TCATTCATAAAGCAATGCTTTTCTAAAGATATAAAGAGTATGATGATCAATA 854  
DB 528 TCATTCATAAAGCAATGCTTTTCTAAAGATATAAAGAGTATGATGATCAATA 587

QY 855 GTTTAGAAAAAGATGTGATCGGTAGCGGTTTTCAAATTTTCAACACCGAAGTAATTAAT 914  
DB 588 GTTTAGAAAAAGATGTGATCGGTAGCGGTTTTCAAATTTTCAACACCGAAGTAATTAAT 647  
QY 915 ATCCCTTTTATCAAGAACCGGCAAAAGAAACGAAATGATACCAATCACTGTCGAAAAA 974  
DB 648 ATCCCTTTTATCAAGAACCGGCAAAAGAAACGAAATGATACCAATCACTGTCGAAAAA 707  
QY 975 AGATATAATGGAGATAAGACGGTTTCGTTTCGTTTCGTTTCGTTTCGTTTCGTTTCGTT 1034  
DB 708 AGATATAATGGAGATAAGACGGTTTCGTTTCGTTTCGTTTCGTTTCGTTTCGTTTCGTT 767  
QY 1035 CGAAACAGCAAGAAATGGCGGAAACGTTAAAGAGTTATGGAATTAAGACTTAGAAGCAA 1094  
DB 768 CGAAACAGCAAGAAATGGCGGAAACGTTAAAGAGTTATGGAATTAAGACTTAGAAGCAA 827  
QY 1095 ACTTAAGAGTGTGTTGATAGTGCAGTATCTTAAATTTTGTATATAGGAATTTGAAGTTA 1154  
DB 828 ACTTAAGAGTGTGTTGATAGTGCAGTATCTTAAATTTTGTATATAGGAATTTGAAGTTA 887  
QY 1155 AATTAGATGCTTAAAAATTTGTAATTAAGAGAGGAGTATACATGAACAAAAATATTAAT 1214  
DB 888 AATTAGATGCTTAAAAATTTGTAATTAAGAGAGGAGTATACATGAACAAAAATATTAAT 947  
QY 1215 ATTCCTCAAAACCTTTTAAACGAGTGAAGAGTACTCAACCAATATAAATAAATTAATTAAT 1274  
DB 948 ATTCCTCAAAACCTTTTAAACGAGTGAAGAGTACTCAACCAATATAAATAAATTAATTAAT 1007  
QY 1275 TAAAAAGAAACCGATACCGTTTACGAATTTGAAACAGGTAAAGGGGCAATTAACGAGAAAC 1334  
DB 1008 TAAAAAGAAACCGATACCGTTTACGAATTTGAAACAGGTAAAGGGGCAATTAACGAGAAAC 1067  
QY 1335 TGGCTTAAATTAAGTAAACAGGTAAAGTCTATTAATTAAGAGAGTATCTTCACTTAT 1394  
DB 1068 TGGCTTAAATTAAGTAAACAGGTAAAGTCTATTAATTAAGAGAGTATCTTCACTTAT 1127  
QY 1395 CGTCAGAAAAATTAATACTCGTGTACCTTTAATTTCAACAGATATTTTACAGT 1454  
DB 1128 CGTCAGAAAAATTAATACTCGTGTACCTTTAATTTCAACAGATATTTTACAGT 1187  
QY 1455 TTCAATTCCTTAAACAGAGGTATAAATTTGTTGGAGTATTCCTTACCATTTAAGCA 1514  
DB 1188 TTCAATTCCTTAAACAGAGGTATAAATTTGTTGGAGTATTCCTTACCATTTAAGCA 1247  
QY 1515 CACAAATTTAAAAAGTGGTTTTTGAAGCCATCGCTCTGACATCTATCTGATTTGTG 1574  
DB 1248 CACAAATTTAAAAAGTGGTTTTTGAAGCCATCGCTCTGACATCTATCTGATTTGTG 1307  
QY 1575 AAGAAGGATTTTCAAGCGTACCTTTGGATATTCACGAAACCTAGGGTGTCTTGGCA 1634  
DB 1308 AAGAAGGATTTTCAAGCGTACCTTTGGATATTCACGAAACCTAGGGTGTCTTGGCA 1367  
QY 1635 CTCAGTCTCGATTCAGCAATTCCTTAAGCTGCGAGCGGAATGCTTTTCATCTCTAAACCA 1694  
DB 1368 CTCAGTCTCGATTCAGCAATTCCTTAAGCTGCGAGCGGAATGCTTTTCATCTCTAAACCA 1427  
QY 1695 AAGTAAACAGTGTCTTAAATAAATTTACCGGCATACCAAGATGTTCCAGATAAATATT 1754  
DB 1428 AAGTAAACAGTGTCTTAAATAAATTTACCGGCATACCAAGATGTTCCAGATAAATATT 1487  
QY 1755 GGAAGCTATATAGTACTTGTTCAAATTTGTTCAATCGGATATCTGTAACCTGTTTA 1814  
DB 1488 GGAAGCTATATAGTACTTGTTCAAATTTGTTCAATCGGATATCTGTAACCTGTTTA 1547  
QY 1815 CTAAAAATCAGTTTTCATCAAGCAATGAACACGCGCAAGTAAACAAATTAAGTACCGTTA 1874  
DB 1548 CTAAAAATCAGTTTTCATCAAGCAATGAACACGCGCAAGTAAACAAATTAAGTACCGTTA 1607  
QY 1875 CTTATGAGCAAGTATGCTATTTTAAATTAAGTATCTATTTTAAACCGGAGGAATTAAT 1934  
DB 1608 CTTATGAGCAAGTATGCTATTTTAAATTAAGTATCTATTTTAAACCGGAGGAATTAAT 1667  
QY 1935 TCTATGAGTGGCTTTTGTAAAAATTTGAAAAATTTACAGTTTCTTAAAGGGAATGTAGATAA 1994

Db 1668 TCTATGAGTCGCTTTTGTAAATTTGGAAAGTTACACGTTACTAAAGGGAATGTAGATAAA 1727  
Qy 1995 TTATAGGTATCTACTGACAGCTTCCAGAGCTTAAAGAGTCCCTAGCGCTCTTATCA 2054  
Db 1728 TTATAGGTATCTACTGACAGCTTCCAGAGCTTAAAGAGTCCCTAGCGCTCTTATCA 1787  
Qy 2055 TGGGGAAGCTCGGATCATATGCAAGACAAATAAATAACTCGCAACAGACACTTGGAGAAATGG 2114  
Db 1788 TGGGGAAGCTCGGATCATATGCAAGACAAATAAATAACTCGCAACAGACACTTGGAGAAATGG 1847  
Qy 2115 GACGAATCGAGAAACCCCTTTTACGCTGATACATATCTAATAAAGCCGTAAAGAGAC 2174  
Db 1848 GACGAATCGAGAAACCCCTTTTACGCTGATACATATCTAATAAAGCCGTAAAGAGAC 1907  
Qy 2175 GGGTTCAAAAGGTTTAAATAAAGGAGAACCAATCAATGCAATTAGCTAGAACTATATTTT 2234  
Db 1908 GGGTTCAAAAGGTTTAAATAAAGGAGAACCAATCAATGCAATTAGCTAGAACTATATTTT 1967  
Qy 2235 TTGACAACTGGAGAAATTTAGAGAACGCTCTCTCCAGACCAAGTTTACAAAGAGCTAGTG 2294  
Db 1968 TTGACAACTGGAGAAATTTAGAGAACGCTCTCTCCAGACCAAGTTTACAAAGAGCTAGTG 2027  
Qy 2295 CACTTAAACATAATTTAATTAACGCTATAGTGTGGAGACACTGTATATATGGRABAGCCG 2354  
Db 2028 CACTTAAACATAATTTAATTAACGCTATAGTGTGGAGACACTGTATATATGGRABAGCCG 2087  
Qy 2355 TAGAAGAAATTTAAAGCAAGAGAGAAATTTAGAGAGAAATTTAATGCCATATGCGTGGCCGT 2414  
Db 2088 TAGAAGAAATTTAAAGCAAGAGAGAAATTTAGAGAGAAATTTAATGCCATATGCGTGGCCGT 2147  
Qy 2415 TAGGATGGGACATATCAATTTCTTGGAGAAATACAAATTTGAAGGATACATGACACTG 2474  
Db 2148 TAGGATGGGACATATCAATTTCTTGGAGAAATACAAATTTGAAGGATACATGACACTG 2207  
Qy 2475 GGCAAAATGAATTTACGCTCTTACGTATATAAGAGCCGTTTATTTCTTAATATAACGGCT 2534  
Db 2208 GGCAAAATGAATTTACGCTCTTACGTATATAAGAGCCGTTTATTTCTTAATATAACGGCT 2267  
Qy 2535 CTTTTTATAGAAAATCTTACGCTGTTTTCGAAATCTGCGGTACCCCAAG 2594  
Db 2268 CTTTTTATAGAAAATCTTACGCTGTTTTCGAAATCTGCGGTACCCCAAG 2327  
Qy 2595 ATTAGAAATCAGTAGATCAAAATTTTACCAATAGAAATCAGGAAATCAGATCCAAACCAT 2654  
Db 2328 ATTAGAAATCAGTAGATCAAAATTTTACCAATAGAAATCAGGAAATCAGATCCAAACCAT 2387  
Qy 2655 AAAAACTAGAACAAATTCGAAAGTTAATACTCAAGCTAGTAGTGGAATTAATCCC 2714  
Db 2388 AAAAACTAGAACAAATTCGAAAGTTAATACTCAAGCTAGTAGTGGAATTAATCCC 2447  
Qy 2715 AATGAGCCACAGAACCCAGGACCAAGAACAGATCAGACCAAGTAACTTGGATTAGA 2774  
Db 2448 AATGAGCCACAGAACCCAGGACCAAGAACAGATCAGACCAAGTAACTTGGATTAGA 2507  
Qy 2775 AATGGAAGAGAAAAAGCAATGACTTCGTGTGAATAATGCAGAAATCGTTGCTTATTT 2834  
Db 2508 AATGGAAGAGAAAAAGCAATGACTTCGTGTGAATAATGCAGAAATCGTTGCTTATTT 2567  
Qy 2835 TTTTTTAAAGCGGTATCTAGATATAACGAAACAGCAATCAATAGAAACGAAAG 2894  
Db 2568 TTTTTTAAAGCGGTATCTAGATATAACGAAACAGCAATCAATAGAAACGAAAG 2627  
Qy 2895 AGCCATGACACATTTATAAATGTTTGAACACATTTTATAAATGCAATAGCCCGATAGAT 2954  
Db 2628 AGCCATGACACATTTATAAATGTTTGAACACATTTTATAAATGCAATAGCCCGATAGAT 2687  
Qy 2955 TGCCAAACCAAGCTTATAGTAGTACAGATGAATCTTCCCTCGTAAGAGTTATTTAA 3014  
Db 2688 TGCCAAACCAAGCTTATAGTAGTACAGATGAATCTTCCCTCGTAAGAGTTATTTAA 2747  
Qy 3015 TTAACCTTTGTTGAAGACGCTATATACCGTACTATCATATATAGGAAATCAGAGAT 3074

Db 2748 TTAACCTTTGTTGAAGACGCTATATAACCGTACTATCATTTATATAGGAAATCAGAGAT 2807  
Qy 3075 TTTTCAAGTATCTTAAGCTACTGAAATTTTAAAGATTTGTTTAAAGCAATCAATCGGAAATCGTTTG 3134  
Db 2808 TTTTCAAGTATCTTAAGCTACTGAAATTTTAAAGATTTGTTTAAAGCAATCAATCGGAAATCGTTTG 2867  
Qy 3135 ATTTGCTTTTGTATTTCAATTTATAGAAGGTGGAGTTTGTATGATCATGATGAATGATA 3194  
Db 2868 ATTTGCTTTTGTATTTCAATTTATAGAAGGTGGAGTTTGTATGATCATGATGAATGATA 2927  
Qy 3195 AACTTATATAAATAATGTTTATTTGGAGATAGAAATTTAGCAAAATATCTATACACTAGA 3254  
Db 2928 AACTTATATAAATAATGTTTATTTGGAGATAGAAATTTAGCAAAATATCTATACACTAGA 2987  
Qy 3255 AACGTTTAAAGAAAGAGTTAGAAAAAGAGAAATATCTACTTAGAAAAAATAATCAGATAAGTA 3314  
Db 2988 AACGTTTAAAGAAAGAGTTAGAAAAAGAGAAATATCTACTTAGAAAAAATAATCAGATAAGTA 3047  
Qy 3315 TTTTCTTCGGAGGGGGAAGATTTATATATATAAGTTTATAGAAAAATACAAAAATTAATTTA 3374  
Db 3048 TTTTCTTCGGAGGGGGAAGATTTATATATATAAGTTTATAGAAAAATACAAAAATTAATTTA 3107  
Qy 3375 TTTGATTTAGTGGAAAAAATTTGACTTATAAAGGAAAAAATCTTTTCAAAACATGCAAT 3434  
Db 3108 TTTGATTTAGTGGAAAAAATTTGACTTATAAAGGAAAAAATCTTTTCAAAACATGCAAT 3167  
Qy 3435 ATTTGAAACAGTTGAATGAAAAAGCAAAACCAAGTTTAAATTAACCACTATTTTATAGGAT 3494  
Db 3168 ATTTGAAACAGTTGAATGAAAAAGCAAAACCAAGTTTAAATTAACCACTATTTTATAGGAT 3227  
Qy 3495 TATAGAAAGAGAGACGCTGAATGAATATCCCTTTTGTGTAGAAATCTGCTTTCATGA 3554  
Db 3228 TATAGAAAGAGAGACGCTGAATGAATATCCCTTTTGTGTAGAAATCTGCTTTCATGA 3287  
Qy 3555 CGGCTTGTAAAGTACAAATTTAAAAATAGTAAAAATTCGCTCAATCACTACCAAGCCAGG 3614  
Db 3288 CGGCTTGTAAAGTACAAATTTAAAAATAGTAAAAATTCGCTCAATCACTACCAAGCCAGG 3347  
Qy 3615 TAAAGCAAAAGGGCTATTTTTCGCTCAAAATCAAGCATGATGGGGTCTGG 3674  
Db 3348 TAAAGCAAAAGGGCTATTTTTCGCTCAAAATCAAGCATGATGGGGTCTGG 3407  
Qy 3675 TGTGTTCTGACTTCGAGGAGCGATTTCAAGAAAAATCAAGATACATTTACACATTTGAC 3734  
Db 3408 TGTGTTCTGACTTCGAGGAGCGATTTCAAGAAAAATCAAGATACATTTACACATTTGAC 3467  
Qy 3735 ACCCAAGCTTTATCGTTATGGAACGTTATGAGAACCGTTTATACAGAAAGGACA 3794  
Db 3468 ACCCAAGCTTTATCGTTATGGAACGTTATGAGAACCGTTTATACAGAAAGGACA 3527  
Qy 3795 TTTCTGAAAAAATTTAAGACAAATCAATACCTCTTTTATTTGATTTTGATAATTCACACGCG 3854  
Db 3528 TTTCTGAAAAAATTTAAGACAAATCAATACCTCTTTTATTTGATTTTGATAATTCACACGCG 3587  
Qy 3855 AAAAGAAAATTTTCAGCAAGCGATTTTAACTCTGCAAGCCGCTATTTGATTTTATGCTTATGCC 3914  
Db 3588 AAAAGAAAATTTTCAGCAAGCGATTTTAACTCTGCAAGCCGCTATTTGATTTTATGCTTATGCC 3647  
Qy 3915 TACTATGATTTCAATCTGATTAAGGTTATCAAGCATATTTTCTTTTATAGAAACGCGAT 3974  
Db 3648 TACTATGATTTCAATCTGATTAAGGTTATCAAGCATATTTTCTTTTATAGAAACGCGAT 3707  
Qy 3975 CTATGCTCTTCAAAATCAGAAATTTAAATCTGCAAGCCGCTATTTTATGCTTATGCTTATGCCAAA 4034  
Db 3708 CTATGCTCTTCAAAATCAGAAATTTAAATCTGCAAGCCGCTATTTTATGCTTATGCTTATGCCAAA 3767  
Qy 4035 TATCCGAGAAATTTTGGAAAGTCTTTGCCAGTTGATCTAAGCTGATCAATTTTGGTAT 4094  
Db 3768 TATCCGAGAAATTTTGGAAAGTCTTTGCCAGTTGATCTAAGCTGATCAATTTTGGTAT 3827  
Qy 4095 TGCTCGCATACCAAGACCGCAATTTTGTAGATTTTGTCTTAAATACCGTATTTCTTT 4154  
Db 3828 TGCTCGCATACCAAGACCGCAATTTTGTAGATTTTGTCTTAAATACCGTATTTCTTT 3887



QY 4155 CAAGAGTGGCAAGATTGGTCTTTTCAACAAACAGATAAATAAGGGCTTTTACTCGTTCAAG 4214  
DB 3888 CAAGAGTGGCAAGATTGGTCTTTTCAACAAACAGATAAATAAGGGCTTTTACTCGTTCAAG 3947  
QY 4215 TCTAACCGTTTTAAGCGGTACAGAGGCAAAACAAAGTAGATGAACCCCTGGTTTAACTCT 4274  
DB 3948 TCTAACCGTTTTAAGCGGTACAGAGGCAAAACAAAGTAGATGAACCCCTGGTTTAACTCT 4007  
QY 4275 CTTATTGACAGAAAGAAATTTTCAGAGAAAGGGTTTAAATAGGGCGTAAATACGTCAT 4334  
DB 4008 CTTATTGACAGAAAGAAATTTTCAGAGAAAGGGTTTAAATAGGGCGTAAATACGTCAT 4067  
QY 4335 GTTTACCCCTCTCTTTAGCGTACTTTAGTTCAAGGCTATTCAATCGAAACGTCGCAATATAA 4394  
DB 4068 GTTTACCCCTCTCTTTAGCGTACTTTAGTTCAAGGCTATTCAATCGAAACGTCGCAATATAA 4327  
QY 4395 TATGTTGAGTTTAAATAATGATAGATCAACCCCTTAGAGAAAGAAAGATTAATCAAAAT 4454  
DB 4128 TATGTTGAGTTTAAATAATGATAGATCAACCCCTTAGAGAAAGAAAGATTAATCAAAAT 4187  
QY 4455 TGTTAGAGTGCCTATTTCAGAAACTATCAAGGGCTAAATAGGGAATACATTACCATCT 4514  
DB 4188 TGTTAGAGTGCCTATTTCAGAAACTATCAAGGGCTAAATAGGGAATACATTACCATCT 4247  
QY 4515 TTGCAAGCTTTGGGTATCAAGTGATTTAAACAGTAAAGATTTATTTGTCGGTCAAGGGTG 4574  
DB 4248 TTGCAAGCTTTGGGTATCAAGTGATTTAAACAGTAAAGATTTATTTGTCGGTCAAGGGTG 4307  
QY 4575 GTTTAAATCAAGAAAGAAAGCGACGTCACGCTTCATTTGTCAGATGGAAGA 4634  
DB 4308 GTTTAAATCAAGAAAGAAAGCGACGTCACGCTTCATTTGTCAGATGGAAGA 4367  
QY 4635 AGATTAAATGGCTTATATTAGCGAAAGAGCGATGATACAGGCTTTATTATGAGCAC 4694  
DB 4368 AGATTAAATGGCTTATATTAGCGAAAGAGCGATGATACAGGCTTTATTATGAGCAC 4427  
QY 4695 CAAAAAGAGATTAGAAAGTCTAGGATTCCTGAACGACATTAATTAATGCTGAA 4754  
DB 4428 CAAAAAGAGATTAGAAAGTCTAGGATTCCTGAACGACATTAATTAATGCTGAA 4487  
QY 4755 GGTACTGAAGCGAATCAGGAAATTTCTTTAAGATTAAACAGGAAAGAAATGCTGCAT 4814  
DB 4488 GGTACTGAAGCGAATCAGGAAATTTCTTTAAGATTAAACAGGAAAGAAATGCTGCAT 4547  
QY 4815 TCAATCTGATGTTAAATCAATTTGCTATCGATCTTAAGATTAAAGTAAAGAAAGAAAA 4874  
DB 4548 TCAATCTGATGTTAAATCAATTTGCTATCGATCTTAAGATTAAAGTAAAGAAAGAAAA 4607  
QY 4875 AGAAGCTATATAAGGCGCTGACAAATTTCTTTGACTTAGAGCATACATTCAATCAAGA 4934  
DB 4608 AGAAGCTATATAAGGCGCTGACAAATTTCTTTGACTTAGAGCATACATTCAATCAAGA 4667  
QY 4935 GACTTTAAACAGCTAGCAGAACCGCTTAAACCGGACACAACTCGATTTGTTAGCTA 4994  
DB 4668 GACTTTAAACAGCTAGCAGAACCGCTTAAACCGGACACAACTCGATTTGTTAGCTA 4727  
QY 4995 TCATACAGCTGAAATTAACCCGCACTATGCCATTATCTATCTATGATACGCTGT 5054  
DB 4728 TGATACAGCTGAAATTAACCCGCACTATGCCATTATCTATCTATGATACGCTGT 4787  
QY 5055 TTGTTTTTTCTTTGCTTTAGCGAATGATTAGCAGAAATATACAGAGTAAAGATTTTAAAT 5114  
DB 4788 TTGTTTTTTCTTTGCTTTAGCGAATGATTAGCAGAAATATACAGAGTAAAGATTTTAAAT 4847  
QY 5115 TAAATTTAGGGGAGAGGAGAGTAGCCGAAACTTTTGTGCTGGACTGAAC 5174  
DB 4848 TAAATTTAGGGGAGAGGAGAGTAGCCGAAACTTTTGTGCTGGACTGAAC 4907  
QY 5175 GAAAGTGAAGGAAAGGCTACTAAACCGTCAGGGGCGAGTGAGAGCGAAGCGAACTTTGAT 5234  
DB 4908 GAAAGTGAAGGAAAGGCTACTAAACCGTCAGGGGCGAGTGAGAGCGAAGCGAACTTTGAT 4967

QY 5235 TTTTAAATTTTCTATCTTTTATAGGTCATTAGAGTATATCTTTATTTGCTTATAAATCTATT 5294  
DB 4968 TTTTAAATTTTCTATCTTTTATAGGTCATTAGAGTATATCTTTATTTGCTTATAAATCTATT 5027  
QY 5295 TAGAGCATATATAGATTTTATGAATAGTCAATTTAAGTTGAGCATATTTAGAGGAGGAAAA 5354  
DB 5028 TAGAGCATATATAGATTTTATGAATAGTCAATTTAAGTTGAGCATATTTAGAGGAGGAAAA 5087  
QY 5355 TCTTGGAGAAATATTTTGAAGAACCCGATTACATGATTTGGATTAGTTCTTTGTTGTTAGCT 5414  
DB 5088 TCTTGGAGAAATATTTTGAAGAACCCGATTACATGATTTGGATTAGTTCTTTGTTGTTAGCT 5147  
QY 5415 GGTTTTAACTAAAGTAGTAGTGAATTTTGGATTTTGGTGTGTGTCTTTGTTGTTAGTAT 5474  
DB 5148 GGTTTTAACTAAAGTAGTAGTGAATTTTGGATTTTGGTGTGTGTCTTTGTTGTTAGTAT 5207  
QY 5475 TTGCTAGTCAAGTGCATTAAATA 5497  
DB 5208 TTGCTAGTCAAGTGCATTAAATA 5230  
  
RESULT 3  
US-09-100-703A-27  
; Sequence 27, Application US/09100703A  
; Patent No. 6610300  
; GENERAL INFORMATION:  
; APPLICANT: SEGERS, Ruud PAM  
; APPLICANT: WATERFIELD, Nicolas R  
; APPLICANT: FRANDSEN, Peer L  
; APPLICANT: WELLS, Jeremy M.  
; TITLE OF INVENTION: COLISTRIDUM PERFRINGENS VACCINE  
; FILE REFERENCE: 97288 US  
; CURRENT APPLICATION NUMBER: US/09/100,703A  
; PRIOR FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: EP97201888.1  
; PRIOR FILING DATE: 1997-06-20  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 27  
; LENGTH: 5231  
; TYPE: DNA  
; ORGANISM: Clostridium perfringens  
US-09-100-703A-27  
  
Query Match 90.0%; Score 4944.6; DB 4; Length 5231;  
Best Local Similarity 99.6%; Pred No. 0;  
Matches 4975; Conservative 0; Mismatches 4; Indels 18; Gaps 1;  
  
QY 501 TAGTAGATCCGCTGCTAAACAAAGCCCGAAAGAGCTGAGTTGGCTGCTGCCACCGCTG 560  
DB 253 TCGCGATCCGCTGCTAAACAAAGCCCGAAAGAGCTGAGTTGGCTGCTGCCACCGCTG 312  
QY 561 AGCAATACTACATAACCCCTTGGGCTCTTAACGGCTCTTAACGGCTTTTGGCTGA 620  
DB 313 AGCAATACTACATAACCCCTTGGGCTCTTAACGGCTCTTAACGGCTTTTGGCTGA 372  
QY 621 AAGGAGGAATATATATCCGATGACCTGCAGGCAAGCTCTAGAAATCGATACGATTTGAAG 680  
DB 373 AAGGAGGAATATATATCCG-----ACTAGAATCGATACGATTTGAAG 414  
QY 681 TGGCAACAGATAAAAAAGAGCTTTAAATTTGTTGCTGAACCTTTTAAACAAAGCAATA 740  
DB 415 TGGCAACAGATAAAAAAGAGCTTTAAATTTGTTGCTGAACCTTTTAAACAAAGCAATA 474  
QY 741 CAATCATTTCTCGCAACAGATAGCGACAGAGGCGAAAAACATTCCTGCTGATCATTC 800  
DB 475 CAATCATTTCTCGCAACAGATAGCGACAGAGGCGAAAAACATTCCTGCTGATCATTC 534  
QY 801 ATAAAGCAATGCTTTTCTTAAAGTAAAAAGTATAAAGACTATGGAATCAATAGTTAG 860  
DB 535 ATAAAGCAATGCTTTTCTTAAAGTAAAAAGTATAAAGACTATGGAATCAATAGTTAG 594  
QY 861 AAAAGATGTGATCCGTAGCGGTTTTTCAAAAATTTTCAACCCAGGAATGAATACTACCT 920





QY 3081 GTATCTAAGCTACTGAAATTTAAGAAATGTTAAGCAATCAATCGGAAATCGTTTGATGCT 3140  
Db 2815 GTATCTAAGCTACTGAAATTTAAGAAATGTTAAGCAATCAATCGGAAATCGTTTGATGCT 2874  
QY 3141 TTTTGTGTTATTCATTTATAGAAAGTGGAGTTTGTATGAATCATCATCAATCAATGTAATAACTTA 3200  
Db 2875 TTTTGTGTTATTCATTTATAGAAAGTGGAGTTTGTATGAATCATCATCAATGTAATAACTTA 2934  
QY 3201 TATAAAAATAGTTTATTTGGAGATAAGAAAAATTAGCAAAATATCTATACACTAGAAACGTT 3260  
Db 2935 TATAAAAATAGTTTATTTGGAGATAAGAAAAATTAGCAAAATATCTATACACTAGAAACGTT 2994  
QY 3261 TAAGAAAGAGTTAGAAAGAGAAATATCTACTTAGAAACAAATATCAATCAATGTAATAACTTA 3320  
Db 2995 TAAGAAAGAGTTAGAAAGAGAAATATCTACTTAGAAACAAATATCAATCAATGTAATAACTTA 3054  
QY 3321 TTCGAGGCGGAAGATTATATATATATAGTTTATAGAAATATCAAAATATCAATCAATGTAATAACTTA 3380  
Db 3055 TTCGAGGCGGAAGATTATATATATATAGTTTATAGAAATATCAAAATATCAATCAATGTAATAACTTA 3114  
QY 3381 TAGTGGAAAAATTTGACCTTATAGAAAGAAAAATCTTTTCAAAACATGCAATATTGAA 3440  
Db 3115 TAGTGGAAAAATTTGACCTTATAGAAAGAAAAATCTTTTCAAAACATGCAATATTGAA 3174  
QY 3441 ACAGTTGAATGAAAGCAAAACCAAGTTTAAATTAACCACTTATTTATAGGATTTATAGG 3500  
Db 3175 ACAGTTGAATGAAAGCAAAACCAAGTTTAAATTAACCACTTATTTATAGGATTTATAGG 3234  
QY 3501 AAAGGAGAACAGCTGAATGAATATCCCTTTTGTGTAGAAATCTGTCTTATAGGAGCTT 3560  
Db 3235 AAAGGAGAACAGCTGAATGAATATCCCTTTTGTGTAGAAATCTGTCTTATAGGAGCTT 3294  
QY 3561 GTTAAAGTACAAATTTAAATATGTAATTCGCTCAATCACTACCAAGCCAGGTAAAG 3620  
Db 3295 GTTAAAGTACAAATTTAAATATGTAATTCGCTCAATCACTACCAAGCCAGGTAAAG 3354  
QY 3621 CAAAGGGGCTATTTTTCGCTATCGCTCAAAATCAAGCATGATTCGCGGTGCTGTGTGT 3680  
Db 3355 CAAAGGGGCTATTTTTCGCTATCGCTCAAAATCAAGCATGATTCGCGGTGCTGTGTGT 3414  
QY 3681 TCTGACTTCCGAGGAGCGATTCAGAAATCAAGATACATTTACATTCGACACCCAA 3740  
Db 3415 TCTGACTTCCGAGGAGCGATTCAGAAATCAAGATACATTTACATTCGACACCCAA 3474  
QY 3741 CGTTTATCGTTATGGAACGTTATGAGACGAAACCGTTTATACACGAAAGGACATTCGA 3800  
Db 3475 CGTTTATCGTTATGGAACGTTATGAGACGAAACCGTTTATACACGAAAGGACATTCGA 3534  
QY 3801 AAACAATTTAGACAAATCAATACCTCTTTTATGATTTGATTTTACACGCGGAAAGA 3860  
Db 3535 AAACAATTTAGACAAATCAATACCTCTTTTATGATTTGATTTTACACGCGGAAAGA 3594  
QY 3861 AACTATTTTACGACGCGATTTTAAACACCGCTATTTAGTTAGTTTATGCTCTACTAT 3920  
Db 3595 AACTATTTTACGACGCGATTTTAAACACCGCTATTTAGTTAGTTTATGCTCTACTAT 3654  
QY 3921 GATTATCAAAATCTGATAAAGGTTATCAAGCATATTTTGTGTTTGAAGCCAGCTATGT 3980  
Db 3655 GATTATCAAAATCTGATAAAGGTTATCAAGCATATTTTGTGTTTGAAGCCAGCTATGT 3714  
QY 3981 GACTTCAAAATCAGAAATTTAAATCTGCAAGACGCGGAAATATTTTTCGAAATATCCG 4040  
Db 3715 GACTTCAAAATCAGAAATTTAAATCTGCAAGACGCGGAAATATTTTTCGAAATATCCG 3774  
QY 4041 AGAATATTTTGGAAAGTCTTTGCCAGTTGATTAACGTTGATTAATCATTTTGTGTTATGCTCG 4100  
Db 3775 AGAATATTTTGGAAAGTCTTTGCCAGTTGATTAACGTTGATTAATCATTTTGTGTTATGCTCG 3834  
QY 4101 CATACCAAGACGGAATAGATTTTTCGCTTAAATTTTACCGTTATTTCTTTCAAGA 4160  
Db 3835 CATACCAAGACGGAATAGATTTTTCGCTTAAATTTTACCGTTATTTCTTTCAAGA 3894

QY 4161 ATGGCAAGATTGGTCTTTTCAAAACAGATAATAAAGGGCTTTTACTCGTCTCAAGTCTAAC 4220  
Db 3895 ATGGCAAGATTGGTCTTTTCAAAACAGATAATAAAGGGCTTTTACTCGTCTCAAGTCTAAC 3954  
QY 4221 GGTTTTAAAGCGGTACAGAAAGCAAAAAACAAGTAGATGAACCCCTGGTTTAACTCTTATT 4280  
Db 3955 GGTTTTAAAGCGGTACAGAAAGCAAAAAACAAGTAGATGAACCCCTGGTTTAACTCTTATT 4014  
QY 4281 GCACGAAACGAAATTTTTCAGGAGAAAGGGTTTAAATAGGGGTAAATAACGTCATGTTTAC 4340  
Db 4015 GCACGAAACGAAATTTTTCAGGAGAAAGGGTTTAAATAGGGGTAAATAACGTCATGTTTAC 4074  
QY 4341 CCTCTCTTTAGCTACTTTTAGTTTCAAGGCTATTCATTCGAAACGTCGGAATATAATATGTT 4400  
Db 4075 CCTCTCTTTAGCTACTTTTAGTTTCAAGGCTATTCATTCGAAACGTCGGAATATAATATGTT 4134  
QY 4401 TGAGTTTAAATTCGATTCATCAACCTTAGAGAAAGAAAGAAAGTAATCAAAATTTGTAG 4460  
Db 4135 TGAGTTTAAATTCGATTCATCAACCTTAGAGAAAGAAAGTAATCAAAATTTGTAG 4194  
QY 4461 AAGTGCCTATTTCAGAAAACTATCAAGGGCTAAATAGGCAATACATTAACCTCTTTGCAA 4520  
Db 4195 AAGTGCCTATTTCAGAAAACTATCAAGGGCTAAATAGGCAATACATTAACCTCTTTGCAA 4254  
QY 4521 AGCTTGGGTATCAAGTGAATTTAAACAGTAAGATTTTATTTGTCGCTCAAGGGTGGTTTAA 4580  
Db 4255 AGCTTGGGTATCAAGTGAATTTAAACAGTAAGATTTTATTTGTCGCTCAAGGGTGGTTTAA 4314  
QY 4581 ATTCAAGAAAAAAGAAAGCAAGCTCAAGCTGTTTCTATTTGTGAGAAATGGAAGAAAGATT 4640  
Db 4315 ATTCAAGAAAAAAGAAAGCAAGCTCAAGCTGTTTCTATTTGTGAGAAATGGAAGAAAGATT 4374  
QY 4641 AATGCTTATATTACGAAAAAAGCGATGTATACAAAGCTTATTTAGTTCAGCAACCAAAA 4700  
Db 4375 AATGCTTATATTACGAAAAAAGCGATGTATACAAAGCTTATTTAGTTCAGCAACCAAAA 4434  
QY 4701 AGAGATTAGAAAGCTTAGGCTTAGGCTTCTGAAAGGCAATTAGATAATTTGCTGAAGTACT 4760  
Db 4435 AGAGATTAGAAAGCTTAGGCTTAGGCTTCTGAAAGGCAATTAGATAATTTGCTGAAGTACT 4494  
QY 4761 GAAGGCGAATCAGGAAATTTTCTTTAAGATTAAACACGAAAGAAATGTTGGCATTTCACT 4820  
Db 4495 GAAGGCGAATCAGGAAATTTTCTTTAAGATTAAACACGAAAGAAATGTTGGCATTTCACT 4554  
QY 4821 TGTAGTGTAAATCATTTGTTGCTATCGATCAATTAAGATTAAGAAAGAAAGAAAGAAAG 4880  
Db 4555 TGTAGTGTAAATCATTTGTTGCTATCGATCAATTAAGATTAAGAAAGAAAGAAAGAAAG 4614  
QY 4881 CTATATAAAGGCGCTGACAAATTTCTTTTGAATAGGCAATACATTCATTCAGAGACTTT 4940  
Db 4615 CTATATAAAGGCGCTGACAAATTTCTTTTGAATAGGCAATACATTCATTCAGAGACTTT 4674  
QY 4941 AAAAAGCTAGCAGAAAGCCCTTAAACCGGACACAACTCGATTTGTTTAGCTATGATAC 5000  
Db 4675 AAAAAGCTAGCAGAAAGCCCTTAAACCGGACACAACTCGATTTGTTTAGCTATGATAC 4734  
QY 5001 AGGCTGAAATTAACAAACCCGCTATGCTTATATCTATCTATGATACGTTGTTGTTT 5060  
Db 4735 AGGCTGAAATTAACAAACCCGCTATGCTTATATCTATCTATGATACGTTGTTGTTT 4794  
QY 5061 TTTCTTTGCTTTAGCGAATGATTAGCAAAATATACAGAGTAAGATTTAAATTAATTA 5120  
Db 4795 TTTCTTTGCTTTAGCGAATGATTAGCAAAATATACAGAGTAAGATTTAAATTAATTA 4854  
QY 5121 TTAGGGGAGAGAGAGAGTAGCCGAAATCTTTTGTGCTTGGACTGAAACCAAGTG 5180  
Db 4855 TTAGGGGAGAGAGAGAGTAGCCGAAATCTTTTGTGCTTGGACTGAAACCAAGTG 4914  
QY 5181 AGGAAAGGCTACTTAAACGTCGAGGGCAGTGAGAGCGAAACCACTTGAATTTTTTA 5240  
Db 4915 AGGAAAGGCTACTTAAACGTCGAGGGCAGTGAGAGCGAAACCACTTGAATTTTTTA 4974  
QY 5241 ATTTCTATCTTTTATAGCTCATTAGATGATATCTTATTTGCTCTATATACTATTTAGCAG 5300

4975 ATTTCTATCTTTATAGGTCATTAGAGTATCTTATTTGCTCTATAAACTATTAGACAG 5034  
5301 CATAATAGATTTTATGAATAGGTCATTAAAGTTGAGCATATTAAGAGAGGAAAATCTTGG 5360  
5035 CATAATAGATTTTATGAATAGGTCATTAAAGTTGAGCATATTAAGAGAGGAAAATCTTGG 5094  
5361 AGAAATATTTGAAGAACCCGATTTACATGGATGAGTATTTCTTTGGTTACGTTGTTT 5420  
5095 AGAAATATTTGAAGAACCCGATTTACATGGATGAGTATTTCTTTGGTTACGTTGTTT 5154  
5421 TAACTAAAGTACGATGAAATTTTGAATTTTGGTGTGTGTCTGTCTGTGTTAGTATTTGCTA 5480  
5155 TAACTAAAGTACGATGAAATTTTGAATTTTGGTGTGTGTGTCTGTCTGTGTTAGTATTTGCTA 5214  
5481 GTCAAAAGTGATTAATA 5497  
5215 GTCAAAAGTGATTAATA 5231

RESULT 4  
US-09-280-428A-11  
; Sequence 11, Application US/09280428A  
; Patent No. 6495738  
; GENERAL INFORMATION:  
; APPLICANT: Folkerts, Otto  
; APPLICANT: Merlo, Donald J  
; TITLE OF INVENTION: Modification of Fatty Acid Composition in Plants by  
; FILE OF INVENTION: Expression of A Fungal Acyl-CoA Desaturase  
; FILE REFERENCE: 50612  
; CURRENT APPLICATION NUMBER: US/09/280, 428A  
; CURRENT FILING DATE: 1999-03-29  
; PRIOR APPLICATION NUMBER: 60/079840  
; PRIOR FILING DATE: 1998-03-30  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 11  
; LENGTH: 10323  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:pDAB1542

US-09-280-428A-11  
Query Match 20.2%; Score 1111; DB 4; Length 10323;  
Best Local Similarity 100.0%; Pred. No. 4.9e-234; Indels 0; Gaps 0;  
Matches 1111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

936 AAAAGAAAACGAAATGATACACCAATCAGTGCAGAAAAGATATTAATGGAGATAAGAC 995  
5286 AAAAGAAAACGAAATGATACACCAATCAGTGCAGAAAAGATATTAATGGAGATAAGAC 5345  
996 GGTTCGTGTCGTGCTGACTTGCACCATATCATAAAATCGAACAGCAAGATGCGCG 1055  
5346 GGTTCGTGTCGTGCTGACTTGCACCATATCATAAAATCGAACAGCAAGATGCGCG 5405  
1056 AAACGTAAAAAGAGTTATGAAATAGACCTTAGAACCAACTTAAAGAGTGTGTGTATGT 1115  
5406 AAACGTAAAAAGAGTTATGAAATAGACCTTAGAACCAACTTAAAGAGTGTGTGTATGT 5465  
1116 GCAGTATCTTAAATTTTGTATTAATAGGAATGGAATTAATAGATGCTAAAAATTTGT 1175  
5466 GCAGTATCTTAAATTTTGTATTAATAGGAATGGAATTAATAGATGCTAAAAATTTGT 5525  
1176 AATTAAGAGAGGTGATTACATGAACAAAATATAAATATCTCAAACTTTTAAACGA 1235  
5526 AATTAAGAGAGGTGATTACATGAACAAAATATAAATATCTCAAACTTTTAAACGA 5585  
1236 GTGAAAAAGTACTCAACCAATATAAACAATTAATTAAGAAAAACCGATACCGTTT 1295  
5586 GTGAAAAAGTACTCAACCAATATAAACAATTAATTAAGAAAAACCGATACCGTTT 5645  
1296 ACGAATTTGGAACAGGTAAAGGCGCAATTACAGCAGAACTGGCTAAAATAGTAAACAGG 1355

5646 ACGAATTTGGAACAGGTAAAGGCGCAATTACAGGAACTCGCTAAAATAAGTAAACAGG 5705  
1356 TAACTGCTTATTAAGTTAGACAGTATCTTATTAATCACTTATCTCAGAAAAATTAACACTGA 1415  
5706 TAACTGCTTATTAAGTTAGACAGTATCTTATTAATCACTTATCTCAGAAAAATTAACACTGA 5765  
1416 ATACTCGTGTCACTTTAATCCCAAGATATTTACAGATTTCATTTCCCTAACAAACAGA 1475  
5766 ATACTCGTGTCACTTTAATCCCAAGATATTTACAGATTTCATTTCCCTAACAAACAGA 5825  
1476 GGTATTAATTTGTTGGAGTATCTTACCAATTTTAAACACCAATATTAATAAAGTGG 1535  
5826 GGTATTAATTTGTTGGAGTATCTTACCAATTTTAAACACCAATATTAATAAAGTGG 5885  
1536 TTTTGAAGGCGCATCGCTCTGACATCTATCTGATTTGTTGAAGAGGATTTCTACAAGCGTA 1595  
5886 TTTTGAAGGCGCATCGCTCTGACATCTATCTGATTTGTTGAAGAGGATTTCTACAAGCGTA 5945  
1596 CTTGGATATTTACCGAACACTAGGTTGCTCTTGCACACTCAAGTCTCGATTACAGCAAT 1655  
5946 CTTGGATATTTACCGAACACTAGGTTGCTCTTGCACACTCAAGTCTCGATTACAGCAAT 6005  
1656 TGCCTAAGCTGCCAGCGGAATGCTTTTATCTCTAAACCAAAAGTAAACAGTCTCTTAATAA 1715  
6006 TGCCTAAGCTGCCAGCGGAATGCTTTTATCTCTAAACCAAAAGTAAACAGTCTCTTAATAA 6065  
1716 AACTTACCGCATACACAGATGTTCCAGATAAATTTGAAGCTATATACGTAATTTG 1775  
6066 AACTTACCGCATACACAGATGTTCCAGATAAATTTGAAGCTATATACGTAATTTG 6125  
1776 TTTCAAAATGGTCAATCGAGAATATCGTCAACTGTTTACTTAAATAACAGTTTCATCAAG 1835  
6126 TTTCAAAATGGTCAATCGAGAATATCGTCAACTGTTTACTTAAATAACAGTTTCATCAAG 6185  
1836 CAATGAAACACGCCAAAGTAAACAAATTTAAGTACCGTTACTTATGAGCAAGTATTTGCTA 1895  
6186 CAATGAAACACGCCAAAGTAAACAAATTTAAGTACCGTTACTTATGAGCAAGTATTTGCTA 6245  
1896 TTTTAAATAGTATCTATTAATTAACCGGAGGAAATATCTATGAGTCTCGCTTTGTAAA 1955  
6246 TTTTAAATAGTATCTATTAATTAACCGGAGGAAATATCTATGAGTCTCGCTTTGTAAA 6305  
1956 TTTGAAAGTTTACACGTTTACTTAAAGGGAATTAGATAAATTTATAGGTATACTACTGACA 2015  
6306 TTTGAAAGTTTACACGTTTACTTAAAGGGAATTAGATAAATTTATAGGTATACTACTGACA 6365  
2016 GCTTCCAAAGGAGCTTAAAGAGTCCCTAGCGC 2046  
6366 GCTTCCAAAGGAGCTTAAAGAGTCCCTAGCGC 6396

RESULT 5  
US-09-056-075-1/c  
; Sequence 1, Application US/09056075  
; Patent No. 5955368  
; GENERAL INFORMATION:  
; APPLICANT: Johnson, Eric A.  
; APPLICANT: Bradshaw, Marite  
; APPLICANT: Rood, Julian  
; TITLE OF INVENTION: Expression System for Clostridium  
; TITLE OF INVENTION: Species  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Quarles & Brady  
; STREET: 1 South Pinckney Street  
; CITY: Madison  
; STATE: WI  
; COUNTRY: US  
; ZIP: 53701-2113  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible





```

; LENGTH: 1167
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-3296

Query Match      5.0%; Score 273.2; DB 4; Length 1167;
Best Local Similarity 55.0%; Pred. No. 1.1e-50;
Matches 630; Conservative 0; Mismatches 498; Indels 18; Gaps 4;

QY 3517 ATGAATATCCCTTTTGGTGTGAAACTGTGCTTCATGACGGCTTGTGTTAAAGTACAAATTT 3576
DB 13 ATGAAAGTTTCTGACGTTATTTTACGTTATTTTAAAAAGTGCCATTCGGTAATACAAGTAT 72
QY 3577 AAAAATAGTAAATTCGCTCAATCACTACCAAGCCAGGTAAAGCAAGAGGGCTATTTTT 3636
DB 73 TTAACAGCAAAATTAACCAATCCGCTATAAGCTGAAGCGGAAAGAGGTGCTATTTTT 132
QY 3637 GCGTATCGCTCAAAATCAAGCATGATGGCGTCTGGTGTGTTCTGACATTCGAGGAA 3696
DB 133 GGGTTAGATCAAGAGGTATGACGGAATCTGTGGAATCGTCTTACTTCTCAAGAA 192
QY 3697 GCGATTCAGAAATCAAGATACATTTACATTTGACACACCAACCGTTATTCGTTATGGA 3756
DB 193 GCTTAGCTGAAATGAAGATAAATTTACTCATTTGACCCCTAATGTTTATAGTTATGCT 252
QY 3757 AGCTATGCAGAGAAACCGTTTCATACAGAAAGGACATCTGAAACAATTTTAAGACAA 3816
DB 253 GCATATGCTGATGATAATAGAACTATGTTAAAGGCCCAACAGAAAGAACTTCCACAA 312
QY 3817 ATCAATACCTCTCTTATTGATTTTGATATTCACACGGCAAGAAATTAATTCAGCAAGC 3876
DB 313 ATCAATACCTCTTGAATCGATTTTGA---TCGCTCACTAGTGGAATATAGATTCTCAA 369
QY 3877 GATATTTAAACCGCTATTTGATTTAGTTTATGCTTACTATGATATCAATCTGAT 3936
DB 370 ATGATCTCGATGCTGCAATTCGAACTGATGCAACGCTAAATTTTAGAAATCTCA 429
QY 3937 AAAGGTTATCAAGCATATTTTCTTTAGAACGCCAGTCTATGTGACTTCAAAATCAGAA 3996
DB 430 GGGGATTTCAAGCGTATTTCTTCTTGAATGCTGCTGATATTTCTTCAAAAATAT 489
QY 3997 TTTAAATCTGTCAAAGCAGCCAAAATAATTCGCAAAATATCCGAGATATTTTGGAAAG 4056
DB 490 TATCAATCAATCGAAGTGCCCAAGAGAGTATCTGAAATTTTAAAGAAAGCAATTTGCTCAA 549
QY 4057 TCTTTGC---CAGTTGATCTACAGTGTATCATTTTGTGTTATGCTCGCATACCAAGAACG 4113
DB 550 GTATTTGCCCTTCGGTCGATTTGGGGTGCACCAATTTTGGTATTTGACGATTTCTTAGAAC 609
QY 4114 GACAAATGTAGATTTTTTTGATCCTTAATTTACCGTTATTTCTTTCAAAGAAATGCGAAGTTGG 4173
DB 610 GACAATGTAGTCTTATCTACTACCAGCACTGACACATGATATGCAACAACTAATCCAGTGG 669
QY 4174 TCTTTCAAACAAACAGATATTAAGGGCTTTACTCGTTCAAGTCTTAAACGGTTTAAAGCGGT 4233
DB 670 TCCATGAATTTTGAAGCCAAAAGAAACATATAAAACCGCAATCTTAATGTT-----GTT 723
QY 4234 ACAGAAAGGCAAAAACAGATAGATGAACCTCGGTTTAATCTCTTATTGCAAGAAACGAAA 4293
DB 724 GCTGTAAGATGATGATCAAGATTAAGAGCAGTGGGTTAAAGTTTAAATGTATAATCCGAA 783
QY 4294 TTTTCAGAGAAAGAGGTTTAAATAGGGCGTAAATACGTCATGTTTACCTCTCTTTAGCC 4353
DB 784 GTCGTCGGTACTTAAAGGAAAACTAGGAGCAAACTAGTAAATTTTACTTTTATCTTTGGCT 843
QY 4354 TACTTTAGTTTACGGCTATTTCAATCGAAACGTCGGAATATAATGTTTGTAGTTTAAATAT 4413
DB 844 TACTATGCTCAAAATTTGAAACAGACCGCTGTTTGTATGATATGATGATATTCACAGT 903
QY 4414 CGATTAGATCAACCCCTTAGAGAAAGAAAGATTAATCAAAATTTGTAGAGTGCCTATTCA 4473
DB 904 TCTTTGAAATTAATCTCTATCAGTTTTCAGAAATAAAGAGAAATTTTAAAGAGTGCCTACTCA 963

```



CORRESPONDENCE ADDRESS:  
ADDRESSEE: Christie Parker & Hale, LLP  
STREET: 350 West Colorado Boulevard, Suite 500  
CITY: Pasadena  
STATE: California  
COUNTRY: United States  
ZIP: 91105

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/992,334  
FILING DATE: 17-DEC-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/302,752  
FILING DATE: 24-DEC-1994

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/FR93/00248  
FILING DATE: 12-MAR-1993

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 992/03034  
FILING DATE: 13-MAR-1992

ATTORNEY/AGENT INFORMATION:  
NAME: Prout, D. Bruce

REGISTRATION NUMBER: 20958

REFERENCE/DOCKET NUMBER: C93:31779

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (626) 795-9900

TELEFAX: (626) 577-8800

INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:

LENGTH: 3792 base pairs  
TYPE: nucleic acid

STRANDEDNESS: both  
TOPOLOGY: circular

MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO

ANTI-SENSE: YES  
IMMEDIATE SOURCE:

CLONE: pg+host4  
US-08-992-334-1

Query Match 4.4%; Score 244.2; DB 2; Length 3792;

Best Local Similarity 58.0%; Pred. No. 3.6e-44; Indels 0; Gaps 0;  
Matches 432; Conservative 0; Mismatches 313;

QY 1176 AATTAGAAGGAGTGAATCATGAACAAATAATAAATATCTCAAACTTTTAAACGA 1235

DB 815 AATTAGAAGGAGTGAATCATGAACAAATAATAAATATCTCAAACTTTTAACTT 756

QY 1236 GTGAAAAGTACTCAACCAATAATAAAGATTTAAAGAAACCGATACCGTTT 1295

DB 755 CAAAACATAATATAGATAAATAATGACAAATATAGATTAAATGACATGATAATCT 696

QY 1296 ACGAATTTGAACAGGTAAAGGGGATTTAAGCAGCAACTGGCTAAATAAGTAAACAGG 1355

DB 695 TTGAATCGGCTAGGAAGAGGCAATTTACCTTGNATTAGTAAGAGGTGTAATTCG 636

QY 1356 TAACGCTATTGAATTAGACAGTCACTATCTCACTATCTCGTCAGAAAAATTTAAACGA 1415

DB 635 TAACGCTATTGAATTAGACAGTCACTATCTCACTATCTCGTCAGAAAAATTTAAACGA 576

QY 1416 ATACTCGTGCACCTTTAAATTCACCAAGATATCTCAAGTTTCAATCCCTCAACACAGA 1475

DB 575 ACGTAATTTCCAAAGTTTAAACAAGGATATATGCGAGTTTAAATTTCCCTAAACCAAT 516

QY 1476 GGTATAAAATTTGGGAGTATCTCTTACCATTTAAGCAGCACAATTTATTTAAAAAGTGG 1535

DB 515 CCTATAAAATATATGGTAATATACCTTTAATACATAAGTACGGATATAATACGCAAAATG 456

QY 1536 TTTTGAAGCCATGCGTCTGACATCTATCTGATTGTTGAAGAGGATTTCTACAAGCTA 1595

DB 455 TTTTGTAGTAGTATAGCTAATGAGATTATTTAATCGTGAATACGGGTTTGTCTAAAGAT 396

QY 1596 CTTTGGATATTTCACGAACACTAGGGTTGCTCTTCCACACTCAAGTCTCGATTTCAGCAAT 1655

DB 395 TATTAAATACAAACGCTCATTTGGCAATTTAATGGCAGAGTTGATATTTCTATAT 336

QY 1656 TGTCTTAAGCTGCCAGCGGAATGCTTTTCATCTCTTAAACCAAAAGTAAACAGTGTCTTAAATA 1715

DB 335 TAAAGTATGTTTCCAGAGAAATATTTTCATCTCTTAAACCTTAAAGTGAATAGCTCACATCA 276

QY 1716 AACTTACCCGCCATACACAGATGTTCCAGATAAATATTGGAAGCTATATAGTACTTTG 1775

DB 275 GATTAAAGTAGAAAAAATCAAGAATATCACACAAAGATAAAACAAAGTATAATTTTCG 216

QY 1776 TTTCAAAATGGTCAATCGAGAATATCGTCAACTGTTTACTTAAATAATCAGTTTTCATCAAG 1835

DB 215 TTATGAATGGTTTAAACAAAGATACAGAAATATTTTACAAATAATCAATTTAACAAT 156

QY 1836 CAATGAAACACGCCAAAGTAAACAAATTTAAGTACCGTTTACTTTATGAGCAAGTATTGTCTA 1895

DB 155 CCTTAAACATGACGGAATTTGACGATTTTAAACAATATTAGCTTTTGAACAATTTCTTATCTC 96

QY 1896 TTTTAAATAGTTATCTATTATTAA 1920

DB 95 TTTTCAATAGCTATAAATTTTAA 71

## RESULT 11

US-08-302-752-1/c

; Sequence 1, Application US/08302752

; Patent No. 6025190

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: THERMOSENSIBLE PLASMID

; NUMBER OF SEQUENCES: 3

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/302,752

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: FR 9203034

; FILING DATE: 13-MAR-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: WO FR/93/00248

; FILING DATE: 12-MAR-1993

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3792 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

US-08-302-752-1

Query Match 4.4%; Score 244.2; DB 3; Length 3792;

Best Local Similarity 58.0%; Pred. No. 3.6e-44;

Matches 432; Conservative 0; Mismatches 313; Indels 0; Gaps 0;

QY 1176 AATTAGAAGGAGTGAATCATGAACAAATAATAAATATCTCAAACTTTTAAACGA 1235

DB 815 AATTAGAAGGAGTGAATCATGAACGAGAAATAATAAATATCTCAAACTTTTAACTT 756

QY 1236 GTGAAAAGTACTCAACCAATAATAAATTTAAACCAATTTTAAAAAGGATCCGTTT 1295

DB 755 CAAAACATAATATAGATAAATAATGACAAATATAGATTAATGAAATGAAATGAAATATCT 696

QY 1296 ACAGAAATTTGGAACAGGTAAAGGGCATTTTAACAGCAAGTGGCTAAATAAGTAACAGG 1355  
DB 695 TTGAATCGGCTCAGAAAGGCCATTTTACCCTGAAATAGTAAGAGGTGTAATTCG 636  
QY 1356 TAACGTCTATTGAATAGACAGTCACTATTCACTTATCGTCAGAAAAATTAACACTGA 1415  
DB 635 TAACGTCCATTGAATAGACCATAAATTTATGCAAACTACAGAAAAATAACTTGTTCATC 576  
QY 1416 ATACTCGTGCACATTTAATTCACCAAGNATCTTACAGTTTCAATTCCTTAAGCAAGAGA 1475  
DB 575 ACAGATAATTTCAAGTTTAAACCAAGGATATATTGCGATTGTAATTTCCCTAAACCAAT 516  
QY 1476 GGTATAAAATTTGGGAGTATTCCTTACCAATTTAAGCACACAAATTTAATAAAAGTGG 1535  
DB 515 CCTATAAATATATGGTAATATACCTTATACATAAGTACGGATATAATACGCAAAATG 456  
QY 1536 TTTTGAAGCGATCGCTGCATCTATCTGATGTTGAAGAGGATCTACAGCGTA 1595  
DB 455 TTTTGTAGTATAGCTAATCAGATTTATTTAATCGTGAATACGGGTTTGTAAAGAT 396  
QY 1596 CTTGGATATTACCGAACACTAGGTTGCTCTTGCACTCAAGTCTCGATTCAGCAAT 1655  
DB 395 TATTAATACAAACGCTCATTTGCAATTTTAAATGCGAGAGTTGATATTTCTATAT 336  
QY 1656 TCGTTAAGCTCGAGGGAATGCTTTTCATCTTAACCAAGTAAGTAACAGTGTCTTAATA 1715  
DB 335 TAAGTATGTTCCAGAGATATTTTTCATCTTAAACCTTAAAGTGAATAGCTCACTATCA 276  
QY 1716 AACTTACCGCATACACAGATGTTCCAGATAAATTTGAAGCTATATACGTTTGTG 1775  
DB 275 GATTAAGTAGAAAAATCAAGATATCACAAAGTAATAACAAAGTATTAATTTTCG 216  
QY 1776 TTTCAAAATGGTCAATCGAGATATCGTCACTGTTTACTTAAAAATCAGTTTCAATCAAG 1835  
DB 215 TTATGAATGGTTTAAACAAAGTATACAGAAATATTTTACAAAAATCAATTTAACAAT 156  
QY 1836 CAATGAACAGCCAAAGTAACAAATTTAAGTACGGTCTTATGACCAAGTATTGTCTA 1895  
DB 155 CTTTAAACATGCGAGGAATGACGATTTAAACAAATTTAGCTTTGAACAAATTTCTATCTC 96  
QY 1896 TTTTAAATAGTATCTATTATTTAA 1920  
DB 95 TTTTCAATAGTATAAATTTAA 71

## RESULT 12

US-08-992-334-2/c  
; Sequence 2, Application US/08992334  
; Patent No. 5919678  
; GENERAL INFORMATION:  
; APPLICANT: Gruss, Alexandra  
; TITLE OF INVENTION: METHODS FOR USING A TEMPERATURE-SENSITIVE  
; TITLE OF INVENTION: METHODS FOR USING A TEMPERATURE-SENSITIVE  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Christie Parker & Hale, LLP  
; STREET: 350 West Colorado Boulevard, Suite 500  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: United States  
; ZIP: 91105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/992,334  
; FILING DATE: 17-DEC-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/302,752

; FILING DATE: 24-DEC-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/FR93/00248  
; FILING DATE: 12-MAR-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 992/03034  
; FILING DATE: 13-MAR-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Prout, D. Bruce  
; REGISTRATION NUMBER: 20958  
; REFERENCE/DOCKET NUMBER: C93:31779  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (626) 795-9900  
; TELEFAX: (626) 577-8800  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5234 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: both  
; TOPOLOGY: circular  
; MOLECULE TYPE: DNA (genomic)  
; US-08-992-334-2  
Query Match 4.4%; Score 244.2; DB 2; Length 5234;  
Best Local Similarity 58.0%; Pred. No. 4e-44; Mismatches 313; Indels 0; Gaps 0;  
Matches 432; Conservative 0;  
QY 1176 AATTAAAGAGGAGTACATGAAACAAAAATATAAAATATTCTCAAACTTTTAAACA 1235  
DB 2257 AATTAAGAGGGTTATAATGAACGAGAAAAATATAAAACACAGTCAAACTTTTACTT 2198  
QY 1236 GTGAAAAAGTACTCAACCAATATAAAACAAATTCGAATTTAAAGAAACCGATACCGTTT 1295  
DB 2197 CAAAACATAATATAGATAAAATATGCAAAATATAAGATTAATGAACATGATATATCT 2138  
QY 1296 ACGAAATTTGGAACAGGTAAAGGGCATTTAAACGACAAAACCTGGCTAAAAATAAGTAAACAGG 1355  
DB 2137 TTGAATCGGCTCAGAAAAAGGCCATTTTACCTTGAATTAGTAAGAGGTGTAATTTTCG 2078  
QY 1356 TAACGTCTATTGAATAGACAGTCACTATTCACTTATCGTCAGAAAAATTAACACTGA 1415  
DB 2077 TAACGTCCATTGAATAGACCATAAATTTATGCAAACTTACAGAAAAATTAACCTTGTGATC 2018  
QY 1416 ATACTCGTGCATTTAATTTCAACAGATATTTCAAGTTCCTCAATTCCTCAACAAACAGA 1475  
DB 2017 ACGATAATTTCCAGTTTAAACAGGATATATTCGAGTTTAAATTTCTTAAACCAAT 1958  
QY 1476 GGTATAAAATTTGGGAGTATTCCTTACCATTAAAGCACACAAATTTAAAGAGTGG 1535  
DB 1957 CCTATAAAATATATGTTAATATACCTTATAACATAAGTACGGATATAATACGCAAAATG 1898  
QY 1536 TTTTGAAGCCATCGCTCAGATCTATCTGATTTGTAAGAGGATTTCTACAGCGTA 1595  
DB 1897 TTTTGAATAGTATAGCTAATGAGATTTTAACTGCGAATACCGGTTTGTGTAAGAT 1838  
QY 1596 CTTTGGATATTCACCGAACACTAGGTTTGTCTTGCACACTCAAGTCTCGATTCAGCAAT 1655  
DB 1837 TATTAATACAAACCGCTCATTTGGCATTACTTTTAAATGCGAGAGTTGATATTTCTATAT 1778  
QY 1656 TCGTTAAGCTGCGAGGGAATGCTTTTCATCTTCACTTAAACCAAGTAAGTACGTTCTTAATA 1715  
DB 1777 TAAGTATGTTCCAGAGATATTTTTCATCTTAAACCTTAAAGTGAATAGCTCACTATCA 1718  
QY 1716 AACTTACCGCCATACCAAGATGTTCCAGATAAATTTGGAAGCTATATACGTACTTTG 1775  
DB 1717 GATTAAGTAGAAAAATCAAGATATCACACAAAGATAAAACAAAGTATAATTTATTCG 1658  
QY 1776 TTTCAAAATGGTCAATCAGAAATATCGTCACTGTTTACTTAAAAATCAGTTTCACTCAAG 1835  
DB 1657 TTATGAATGGTTAAACAGAAATATAAAGAAAAATTTTACAAAAATCAATTTTAACAAT 1598  
QY 1836 CAATGAACACGCGCAAGTAAACAAATTTAAGTACCGTTTACTTATGAGCAAGTATTGTCTA 1895

Db 1597 CCTTAAACATGCGAGGATTCGACGATTAAACATATATAGCTTTGAACAATTCCTTATCTC 1538  
Qy 1896 TTTTAAATAGTTATCTATTATTTAA 1920  
Db 1537 TTTTCAATAGCTATAAATTTATTTAA 1513  
RESULT 13  
US-08-302-752-2/c  
; Sequence 2, Application US/08302752  
; Patent No. 6025190  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: THERMOSENSIBLE PLASMID  
; NUMBER OF SEQUENCES: 3  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/302,752  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 9203034  
; FILING DATE: 13-MAR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO FR/93/00248  
; FILING DATE: 12-MAR-1993  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5234 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; US-08-302-752-2  
Query Match 4.4%; Score 244.2; DB 3; Length 5234;  
Best Local Similarity 58.0%; Pred. No. 4e-44;  
Matches 432; Conservative 0; Mismatches 313; Indels 0; Gaps 0;  
Qy 1176 AATTGAAGAGGTGATTACATGACCAAAAATATATATATTTCTCAAACTTTTAAACGA 1235  
Db 2257 AATTGAAGAGGTGATTACATGACCAAAAATATATATATTTCTCAAACTTTTAACTT 2198  
Qy 1236 GTGAAAAGTACTCAACCAATATATAAACAATTTGAATTTAAAGAAACCGATACCGTTT 1295  
Db 2197 CAAAACATATATAGATAAATAATATGCAATATATAGATTAAATGACATGATATCT 2138  
Qy 1296 ACGAAATTGGAACAGGTAAAGGCGATTAAACGAACTGGCTAAATAAGTAAACAGG 1355  
Db 2137 TTGAAATCGGCTCAGGAAAAGGCCATTTTACCTTGAATTAGTAAAGAGGTGTAATTTCG 2078  
Qy 1356 TAAGTCTATTGAATTAGACAGCATCTATTCTCACTTATCGTCAGAAAATTTAAACTGA 1415  
Db 2077 TAATGCTGATTGAATAGACCAATAATATGCAAACTACAGAAAATAAATCTGTTGATC 2018  
Qy 1416 ATACTCGTGCACCTTTAAATCACAAGATATTTACAGTTTCAATTTCCCTTAAACAAACAGA 1475  
Db 2017 ACGATAATTTCCAGTTTAAACAAGGATATATTGCAAGTTTAAATTTCTTAAACCAAT 1958  
Qy 1476 GGTAATAAATTTGGGAGTATTCCTTACCATTTAAGCACACAAATATATAAAGAGTG 1535  
Db 1957 CCTATAAATAATATGGTAATATACCTTTATAACATAAGTAGCGGATATAATACGCAAAATTG 1898  
Qy 1536 TTTTGAAGCGATCGCTGACATCTATCTGATTTGTTGAAGAGGATTTCAACGCTA 1595  
Db 1897 TTTTGTATAGTATAGCTAATGAGATTTTATTTATCTGGAATACGGGTTGCTAAAGAT 1838  
Qy 1596 CCTTGGATATTCACCGAACACTAGGGTTGCTCTTGGCACATCTCAAGTCTCGATTTCAGCAAT 1655

Db 1837 TATTAAATACAAAACGCTCATTTGGCATTACTTTTAAATGCGAGAGTTGATATTTCTATAT 1778  
Qy 1656 TGTCTTAAGCTGCCAGCGGAATGCTTTTCATCTCTAAACCAAAAGTAAACAGTGTCTTAAATA 1715  
Db 1777 TAAAGTATGTTTCCAAAGAGATATTTTTCATCTCTAAACCTTAAAGTGAATAGCTCAGCTTATCA 1718  
Qy 1716 AACTTTACCGCCATACACAGATGTTCCAGATAAATATTGGAAGCTATATAGCTACTTTG 1775  
Db 1717 GATTAACTAGAAAAAATCAAGATATCACAAAGATAAACAAGATATATATTTTCG 1658  
Qy 1776 TTTCAAAATGGTCAATCGAGAATATGTCACACTGTTTCTTAAATAATCAGTTTCTATCAAG 1835  
Db 1657 TTATGAAATGGTTTAAACAAAGATATACAGAAAAATATTACAAAAATCAATTTTAAACAAT 1598  
Qy 1836 CAATGAAACACGCCAAAGTAAACAAATTTAAGTACCGTTTACTTATGAGCAAGTATTGTCTA 1895  
Db 1597 CCTTAAACATGCGAGGAAATGACGATTTTAAACAATTTAGCTTTGAACAATTTCTTATCTC 1538  
Qy 1896 TTTTAAATAGTTATCTATTATTTAA 1920  
Db 1537 TTTTCAATAGCTATAAATTTTAA 1513  
RESULT 14  
US-09-329-920-1/c  
; Sequence 1, Application US/09329920  
; Patent No. 6326206  
; GENERAL INFORMATION:  
; APPLICANT: Bjornvad, Mads Eskelund  
; APPLICANT: Rasmussen, Michael Dolberg  
; APPLICANT: Jorgensen, Per Lina  
; APPLICANT: Borchert, Torben Vedel  
; APPLICANT: Ehrlich, Dr. Stanislas Dusko  
; TITLE OF INVENTION: In Vivo Recombination  
; FILE REFERENCE: 4833-204-US  
; CURRENT APPLICATION NUMBER: US/09/329,920  
; CURRENT FILING DATE: 1999-08-10  
; PRIOR APPLICATION NUMBER: 60/050,590  
; PRIOR FILING DATE: 1997-06-24  
; PRIOR APPLICATION NUMBER: 1471/96  
; PRIOR FILING DATE: 1996-12-20  
; PRIOR APPLICATION NUMBER: 0592/97  
; PRIOR FILING DATE: 1997-05-23  
; PRIOR APPLICATION NUMBER: 0935/97  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 5313  
; TYPE: DNA  
; ORGANISM: Homologous  
; US-09-329-920-1  
Query Match 4.4%; Score 244.2; DB 4; Length 5313;  
Best Local Similarity 58.0%; Pred. No. 4e-44;  
Matches 432; Conservative 0; Mismatches 313; Indels 0; Gaps 0;  
Qy 1176 AATTGAAGAGGTGATTACATGACCAAAAATATATAAATATTTCTCAAACTTTTAAACGA 1235  
Db 4340 AATTGAAGAGGTGATTACATGACCAAAAATATATAAACACAGTCAAAACTTTTACTT 4281  
Qy 1236 GTGAAAAGTACTCAACCAATATATAAACAATTTGAATTTAAAGAAACCGATACCGTTT 1295  
Db 4280 CAAAACATATATAGATAAATAATGACAAATATAAGTTAAATGAAACATGATATCT 4221  
Qy 1296 ACGAAATTGGAACAGGTAAAGGCGATTTAACGAGCAAACTGGCTAAATAAGTAAACAGG 1355  
Db 4220 TTGAAATCGGCTCAGGAAAAGGCCATTTTACCTTGAATTAGTAAAGAGGTGTAATTTCG 4161  
Qy 1356 TAACTGCTATTGAATTAGACAGTCTATCTATTCACTTATCTGTCAGAAAAATTAACATGA 1415  
Db 4160 TAACTGCGCATTAAGATAGACCAATTAATGCAAAACCTACAGAAATTAATCTGTTGATC 4101

```

QY 1416 ATACTCGTGCACCTTTAATTCACCAAGATATTTCTACAGTTTCAATTCCTTAAACAAACAGA 1475
Db 4100 ACGATAATTTCCAAAGTTTAAACCAAGATATATTTGCGAGTTTAAATTTCTTAAACCAAT 4041
QY 1476 GGTATAAATTTGTTGGGATTTCTTACCAATTTAAGCACACAAATTTTAAABAAAGTGG 1535
Db 4040 CCTATAAATATATGTTAAATATACCTTATACATAAGTACGGATATATACGCAAAATG 3981
QY 1536 TTTTGTAAAGCCATCGCTCTCACATCTATCTGATTGTTGAAGAAAGGATTTCTACAAGCGTA 1595
Db 3980 TTTTGTAGTATAGCTAATCAGATTTATTTAATCGTGGATACGGGTTTGTCTAAAGAT 3921
QY 1596 CTTTGGATATTCACGGAACACTAGGTTGCTCTTGACACTCAAGTCTCGATTCAGCAAT 1655
Db 3920 TATTAAATACAAAACGGCTCATTTGCTTCTTAAATGGCAGAAAGTTGATATTTCTATAT 3861
QY 1656 TGCTTAAGCTCCCGGGAATGCTTTCATCTTAAACCAAAAGTAAACAGTGTCTTAAATA 1715
Db 3860 TAAGTATGTTCCAAAGATATTTTCTAATCTTAAACCTTAAAGTGAATGCTCACTATCA 3801
QY 1716 AACTTACCCGCAATCACACAGATGTTCCAGATAATNTGGAAGCTATATACGTACTTGG 1775
Db 3800 GATTAAAGTAGAAAAAATCAAGATATATCACAAAGATAAACAAAGTAAATATTTTCG 3741
QY 1776 TTTCAAAATGCTCAATCGCAATATCGTCAACTGTTTACTTAAAAATCAGTTTCATCAAG 1835
Db 3740 TTATGAAATGGTTTAAACAAAGATACAGAAATATTTTACAAAATCAATTTAAACAT 3681
QY 1836 CAATGAAACACGCCAAAGTAAACAAATTTAAGTACCGTTTACTTATGCAAGAGTATGCTA 1895
Db 3680 CTTTAAACATCGAGGAATGACGATTTAAACATATTAGCTTTGAACAAATTTCTTATCTC 3621
QY 1896 TTTTAAATAGTATCTATTTAA 1920
Db 3620 TTTTCAATAGCTATAAATTTAA 3596

```

RESULT 15

```

US-08-875-154-2/c
; Sequence 2, Application US/08875154
; Patent No. 5882888
; GENERAL INFORMATION:
; APPLICANT: Jorgensen, Streen Troels
; TITLE OF INVENTION: DNA Integration By Transporation
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5882888 No. 5882888disk of No. 5882888th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/875,154
; FILING DATE: 17-JUL-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4381.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6169 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

```

```

; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "psj2739"
; ORIGINAL SOURCE:
; ORGANISM: psj2739
; US-08-875-154-2

Query Match
Best Local Similarity 58.0%; Pred. No. 4.2e-44;
Matches 432; Conservative 0; Mismatches 313; Indels 0; Gaps 0;

QY 1176 AATTAAAGAGGAGTCAATTCATGAAACAAAAATATAAAAAATTTCTCAAAACCTTTTAAACA 1235
Db 1539 AATTAAAGAGGTTTATATGACGAGAAATATAAACAACAGTCAACAGTCAACCTTTTACTT 1480
QY 1236 GTGAAAAAGTACTCAACCAATATAAACAATTTGAATTTAAAGAAACCGATACCGTTT 1295
Db 1479 CAAAACATAATATAGATAAAATTAATGACAAATATAAGATTAATGAACATGATATATCT 1420
QY 1296 ACGAAATTGGACACAGTAAAGGCAATTTAAACGACGAACTGCTAAATAAGTAAACAGG 1355
Db 1419 TTGAAATCGGCTCAGAAAGGCCAATTTACCTTGAATTTAGTAAAGAGGTGTAATTCG 1360
QY 1356 TTAACGTCTATTGAAATAGACAGTCTATCTATTCAACTTATCGTCAGAAAAATTTAAACACTGA 1415
Db 1359 TAACTGCCATTGAAATAGACCAATAAATTTATGCAAACTTACAGAAAAATAAATCTGTGATC 1300
QY 1416 ATACTCGTGTCACTTTAATTCACCAAGATATTTACAGTTTCAATTCCTTCAACAAACAGA 1475
Db 1299 ACGATAATTTTCCAGTTTAAACAAAGGATATTTGCGAGTTTAAATTTCTTAAACCAAT 1240
QY 1476 GGTATAAATTTGTCGGAGTATTCCTTACCAATTTAAAGCACACAAATTTATTTAAACAGTGG 1535
Db 1239 CCTATAAATATATGTTAATATACCTTATAACATAAGTACGGATATATAACGCAAAATG 1180
QY 1536 TTTTGAAGCCATCGCTGCTGACATCTATCTGATGTTGTTGAAGAGGATTTCTACAGCGTA 1595
Db 1179 TTTTGTAGTATAGCTAATGAGATTTATTAATCGTGGAAATACGGGTTTCTCTAAAGAT 1120
QY 1596 CCTGTGATATTCACCGAACACTAGGTTTGTCTTTCACACTCAAGTCTCGATTTCAGCAAT 1655
Db 1119 TATTAAATACAAACGCTCATTTGGCATTTCTTTAATGCGAGAGTTGATATTTCTATAT 1060
QY 1656 TCTTTAAGCTGCCAGCGAAATGCTTTTCAATCTTAAACCAAAAGTAAACAGTGTCTTAATAA 1715
Db 1059 TAAATATGTTTCCAAAGAGATATTTTCTATCTTAAACCTTAAAGTGAATAGCTCACTTATCA 1000
QY 1716 AACTTACCGCCATACACAGATGTTCCAGATAAATATTGGAAGCTATATACGTACTTTG 1775
Db 999 GATTAAAGTAGAAAAAATCAAGATATCACCAAGATATAACAAAGTATATTTATTTTCG 940
QY 1776 TTTCAAAATGGGTCAATCGAGAATATCGTCAACTGTTTCTTAAACCAATCAGTTTCTATCAAG 1835
Db 939 TTATGAATGGTTTAAACAAAGATTAACAGAAAAATTTTACAAAAATCAATTTAAACAAAT 880
QY 1836 CAATGAACACACCCAAAGTAAACATTTTAAAGTACCGTTTACTTATGAGCAAGTATTTGCTA 1895
Db 879 CTTTAAACATCGAGGAATGACGATTTAAACATATTTAGCTTTGAAACAAATTTCTATCTC 820
QY 1896 TTTTAAATAGTATCTATTTAA 1920
Db 819 TTTTCAATAGCTATAAATTTAA 795

```

Search completed: September 2, 2004, 00:17:01  
Job time : 342 secs